

## **REMARKS**

In response to the Office Action mailed June 2, 2004, Applicants have amended the claims, which when considered with the following remarks, is deemed to place the application in condition for allowance. Applicants have also added new claim 138 to the application to better define subject matter to which Applicants are entitled. Support for claim 138 may be found throughout the specification, e.g., Example 4, page 104, line 16 to end of page, and page 107, lines 4-9. Favorable consideration of all pending claims is respectfully requested.

In the Office Action of June 2, 2004, the Examiner has made final the restriction requirement. Thus, claims 5-6, 18-14, 26-27, 45, 48, 51, 54-78, 82-85, 88-89, 93-94, 102 and 122-137 have been withdrawn from consideration. Claims 1-4, 7-17, 25, 28-44, 46-47, 49-50, 52-53, 79-81, 86-87, 90-92, 95-101 and 103-121 are presently examined.

The inventors' oath has been found defective by the Examiner because the application serial number is incorrect. Applicants respectfully submit that the declaration does indeed contain a typographical error on page 2, where the numeral "1" was inadvertently cut off. Specifically, "Serial No. 10/014,10" is listed rather than "Serial No. 10/014,101." A new inventors' declaration has been forwarded to the inventors (who reside in Germany) for their execution. As soon as the document is received by the undersigned, it will be submitted forthright to the Patent Office.

The abstract has been objected to by the Examiner as being longer than 150 words. The Examiner has also noted that the abstract should be limited to less than 150 words and should be descriptive of the elected invention. In response to the objection,

Applicants have amended the abstract as submitted herewith. Withdrawal of the objection is therefore warranted.

The disclosure has been objected to due an embedded hyperlink and/or other form of browser-executable code contained therein. Specifically, the Examiner has objected to the web sites set forth on page 89, lines 4-6; page 87, line 33; page 83, line 6; and page 90, lines 33-35. Applicants respectfully submit that the websites were disclosed on those pages indicated above in order to properly reference the source of tools e.g., computer programs, used in the experiments leading up to the presently claimed invention. Thus, one skilled in the art, reading the present application would be able to duplicate the experiments described therein, using those computer programs provided at those websites. Applicants do not intend to have the hyperlinks be active links. Further, Applicants respectfully request the Examiner disable the hyperlinks when preparing the text to be loaded onto the USPTO website as an issued patent. *See* MPEP 608.01, “Hyperlinks and Other Forms of Browser-Executable Code in the Specification.” Withdrawal of the objection to the disclosure is therefore respectfully requested.

Claims 2 and 3 have been objected to as allegedly drawn to non-elected material. In response to the objection, Applicants have amended claims 2 and 3 to recite the nucleotide sequence set forth in SEQ ID NO:26 and the corresponding amino acid sequence, as set forth in SEQ ID NO:4. Claim 3 has also been objected to for allegedly not incorporating a Sequence Identifier. As presently amended, claim 3 presently recites the Sequence Identifier “SEQ ID NO:37” immediately after the accession number AC005917. An amended Sequence Listing is also submitted herewith, having the nucleotide sequence of AC005917 as it appeared on the NCBI database as of the filing

date of the present application. Also submitted herewith is a statement under 37 C.F.R. §1.825(a) and (b). Exhibit A provides a printout of the AC005917 nucleotide sequence as it appeared in Genbank between April 10, 2000 and April 18, 2002, (version gi:6598497). Withdrawal of the rejection of claims 2 and 3 is therefore warranted.

Claims 28 and 29 have been objected to as allegedly dependent on non-elected claims. In response to this objection, Applicants have amended claims 28 and 29 so that they no longer depend on non-elected claims. Based on the foregoing remarks and amendments, withdrawal of the objection to claims 2, 3, 28 and 29 is respectfully requested.

The Examiner has acknowledged Applicants' claim for foreign priority based on European Patent Application No. 00870132.8, filed June 16, 2000. The Examiner has requested a certified copy of the patent application as required by 35 U.S.C. 119(b). Applicants thank the Examiner for the reminder and will submit a certified copy as soon as possible.

Claims 79-81 have been rejected under 35 U.S.C. § 112, second paragraph, as allegedly indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. In particular, the Examiner's position is that it is not clear what the metes and bounds of "diagnostic" are. In response to the rejection, and in order to expedite prosecution of this application, claims 79-81 have been amended so that they no longer recite "diagnostic." Withdrawal of the rejection of claims 79-81 is therefore respectfully requested.

Claims 1-4, 7-17, 25, 28-44, 46-47, 49-50, 52-53, 79-81, 86-87, 90-92, 95-101 and 103-121 have been rejected under 35 U.S.C. §112, first paragraph, as allegedly violative of the written description requirement. Applicants respectfully submit that the *proper* test for sufficiency of description in a patent application is whether the disclosure of the application relied upon “reasonably conveys to the artisan that the inventor had possession at that time, of the later claimed subject matter.” *In re Kaslow*, 707 F.2d 1366, 1375, 217 USPQ 1089, 1096 (Fed. Cir. 1983).

Exactly how the specification allows one skilled in the art to recognize that an applicant had possession of the claimed invention is not material. *In re Smith*, 481 F.2d 910, 178 USPQ 279 (CCPA 1973). Typically, an applicant conveys that he or she is in possession of the invention by use of descriptive means such as “words, structures, figures, diagrams, formulas, et., that set forth the claimed invention.” *Lockwood v. American Airlines*, 107 F.3d 1565, 1572, 41 UAPQ2d 1961, 1966 (Fed. Cir. 1997). To comply with the written description requirement, it is not necessary that the application describe the invention *ipsis verbis*. *In re Lukach*, 442 F.2d 967, 169 USPQ 795 (CCPA 1971). What *is* required is that an ordinarily skilled artisan recognize from the disclosure that applicants invented the subject matter of the claims, including the limitations recited therein. *Smith*, 481 F.2d at 915, 178 USPQ at 284. Thus, it has been well settled that the written description requirement of 35 U.S.C. § 112, first paragraph, can be satisfied without express or explicit disclosure of a later-claimed invention. *See e.g. In re Herschler*, 591 F.2d 693, 700, 200 USPQ 711, 717 (CCPA). In that particular case, the court stated: “[t]he claimed subject matter need not be described in haec verba to satisfy the description requirement. It is not necessary that the application describe the claim

limitations exactly, but only so clearly that one having ordinary skill in the pertinent art would recognize from the disclosure that appellants invented processes including those limitations.” (citations omitted). *See also Purdue Pharma L.P. v. Faulding, Inc.*, 230 F.3d 1320, 1323, 56 USPQ2d 1481, 1483 (Fed. Cir. 2000)(“In order to satisfy the written description requirement, the disclosure as originally filed does not have to provide *en haec verba* support for the claimed subject matter at issue.”)

The burden of showing that the claimed invention is not described in the specification rests on the PTO in the first instance, and it is up to the PTO to give reasons why a description not in *ipsis verbis* is insufficient. *In re Wertheim*, 541 F.2d 257, 191 USPQ 90 (CCPA 1976).

The written description requirement has been addressed by the Court of Appeals for the Federal Circuit in the context of DNA-related inventions in *Enzo Biochem. Inc., v. Gen-Probe Inc.*, 296 F.3d 1316, 63 USPQ2d 1609 (Fed. Cir. 2002). In particular, the court adopted the standard that “the written description requirement can be met by ‘showing that an invention is complete by disclosure of *sufficiently detailed, relevant identifying characteristics*, ...i.e., complete or partial structure, other physical and/or chemical properties, functional characteristics when coupled with a known or disclosed correlation between function and structure, or some combination of such characteristics. (Emphasis added). *Id at 1324, 63 USPQ2d at 1613*. The court in *Enzo* adopted its standard from the USPTO’s Written Description Examination Guidelines. *See 296 F.3d at 1324, 63 USPQ2d at 1613* (citing the Guidelines). In addition to nucleotide sequences, the Guidelines also apply to amino acid sequences.

As presently amended, claim 1 recites: “[a] method for producing a plant having enhanced root growth or for producing a plant having enhanced formation of primary, lateral or adventitious roots, said method comprising: (a) introducing into a plant cell or plant part, an isolated nucleic acid molecule encoding a plant cytokinin oxidase wherein the plant cytokinin oxidase comprises the amino acid sequence motifs GHS (SEQ ID NO:38), VGGTLSN (SEQ ID NO: 39), VLGGLGQFG (SEQ ID NO:40), and ITRARI (SEQ ID NO:41), (b) regenerating a plant therefrom; and (c) selecting a plant having enhanced root growth or enhanced formation of primary, lateral, or adventitious roots.”

Support for amendments to claim 1 may be found throughout the specification and in the literature extant at the time of filing the present application. For example, the individual sequence motifs may be discerned from the sequence alignments in Figure 2. Further, the specification teaches at page 5, lines 23-25 that CKX is an FAD binding enzyme. At the time of filing the present application, CKX was known to contain an FAD binding domain. *See* Houba-Herin (1999) *Plant Journal* 17:615-626 (provided herewith as Exhibit B), Morris et al. (1999) *Biochem. Biophys. Res. Comm.* 255:328-333; (provided herewith as Exhibit C), and Rinaldi A., and Comandini O. (1999) *Trends in Plant Science* 4:300, (provided herewith as Exhibit D).

Houba-Herin et al. (1999) teach that a conserved GHS motif exists as a putative FAD-binding site through a histidine residue of many flavoprotein oxidoreductases. On page 619, column 1 of Houba-Herin et al., the authors state that [t]he deduced CKO amino acid sequence shows sequence similarity with a FAD-binding domain (a.a. approximately 170-240) found in several oxidases, suggesting that CKO is a FAD-dependent oxidase (Figure 5). A GHS motif is also found in the CKO sequence (a.a.

104-106). This motif is typical for those enzymes that covalently bind to FAD through a histidine residue.” As presently amended, claim 1 recites both the GHS (SEQ ID NO:38) motif as well as sequence motifs found in the FAD binding domain described by Houba-Herin et al. in the vicinity of amino acid residues 170-240 in Figure 4 of Houba-Herin et al., i.e. VGGTLSN (SEQ ID NO:39), VLGGLGQFC (SEQ ID NO:40), and ITRARI (SEQ ID NO:41).

As presently amended therefore, claim 1 meets the written description requirement because the application discloses and claim 1 presently recites *sufficiently detailed, relevant identifying characteristics*, i.e., partial structure and functional characteristics when coupled with a known correlation between function and structure. *Enzo Biochem. Inc., v. Gen-Probe Inc.*, 296 F.3d 1316, 1324, 63 USPQ2d 1609, 1613.

With respect to the other claims which have been rejected as allegedly not meeting the written description requirement of 35 U.S.C §112, first paragraph, the Examiner has apparently predicated the rejection on *inter alia*, the fact that Applicants’ disclose one cDNA sequence as AtCKX2 (Arabidopsis thaliana cytokinin oxidase-like gene) of SEQ ID NO:26 encoding SEQ ID NO:4 (page 88, 1<sup>st</sup> full paragraph). In response to this, Applicants submit, the specification is replete with written description of sequences falling within Applicants’ invention as presently claimed. *See e.g.*, specification, pages 18-23 and 36-51. Perhaps the Examiner meant to say that SEQ ID NO:26 is particularly exemplified, while other specifically disclosed and claimed sequences are not. Exemplification, however, is not required to fulfill the written description requirement under 35 U.S.C. §112, first paragraph.

On the first point, the tacitly assumed need for exemplification, we do not regard 112, first paragraph, as requiring a specific example of everything within the scope of a broad claim ... What the Patent Office is here apparently attempting is to limit all claims to specific examples, notwithstanding the clear disclosure of a broader invention. This it may not do.

*In re Anderson*, 176 USPQ331 (Bd. Pat. App. 1973)

The Examiner also apparently predicates the rejection on the fact that “[a]pplicants do not identify essential regions of AtCKX2 protein encoded by SEQ ID NO:26, nor do Applicants describe any polynucleotide sequences that hybridize to SEQ ID NO:26 or nucleic acids that encode a functional fragment of SEQ ID NO:4 and retains cytokinin oxidase activity.” Under the proper legal standard for meeting the written description requirement set forth hereinabove, the fact that the specification does not “identify essential regions of AtCKX2 protein encoded by SEQ ID NO:26” (Office Action, page 6), nor specifically exemplify a polynucleotide sequence that hybridizes to SEQ ID NO:26 (*Id*) is not dispositive.

As stated by the Federal Circuit in *Enzo*, “the written description requirement can be met by showing ‘that an invention is complete by disclosure of sufficiently detailed, relevant identifying characteristics...i.e., complete or partial structure, other physical and/or chemical properties, functional characteristics when coupled with a known or disclosed correlation between function and structure, or some combination of such characteristics.’” 296 F.3d at 1324, 63 USPQ2d at 1613.

The present application provides the nucleotide sequence for *Arabidopsis thaliana* AtCKX2 cDNA as set forth in SEQ ID NO:26, as well as an 84 bp fragment corresponding to nucleotides 1171-1254 of the AtCKX2 cDNA as set forth in SEQ ID



NO:31. Amino acid sequences for the full length cDNA as well as the fragment are set forth in SEQ ID NO:4 and 32, respectively. The specification further teaches at pages 18-23 and 36-51, nucleic acid molecules hybridizing to SEQ ID NO:26, nucleic acid molecules that are divergent from a nucleic acid encoding SEQ ID NO:4, nucleic acid molecules that are divergent from those nucleic acids as specified in claim 2 due to differences between alleles, functional fragments of said nucleic acids having biological activity of a cytokinin oxidase, a nucleic acid encoding an amino acid sequence comprising SEQ ID NO:32 and which is at least 70% similar to the amino acid sequence of SEQ ID NO:4, a nucleic acid encoding any immunologically active fragment of a cytokinin oxidase encoded by SEQ ID NO:26 or any immunologically active fragment encoded by any of the nucleic acids specified in claim 3, and a nucleic acid encoding any functional fragment of any cytokinin oxidase encoded by SEQ ID NO:26 or any previously mentioned sequence.

Applicants respectfully submit that Figure 2 provides an alignment of different cytokinin oxidases where the known signature sequences such as GHS, VLGGLGQFC, ITRARI, VGGTLSN and PHPWLN are clearly visible. Those of skill in the art would know that the more conserved a part of a protein is, the more important that part is for structure or function of that protein. Thus, one skilled in the art would know to avoid amino acid substitutions in such areas. Even further, the specification teaches the use of a standard assay for cytokinin oxidase activity based on the conversion of [2-3H] iP to adenine as described in Motyka et al. 1996 *Plant Physiol.* 112:1035-1043 and Example 4, Table 7.

Thus, it is respectfully submitted that the Examiner has failed to establish by a preponderance of the evidence, why one of ordinary skill in the art, who is in possession of the very specific chemical structures of nucleotide sequences that encode a polypeptide of SEQ ID NO:4 and a nucleotide sequence comprising the sequence set forth in SEQ ID NO:26, would be unable to recognize, upon reading the present specification including the alignment of corn and Arabidopsis CKX proteins that Applicants invented the presently claimed subject matter including allelic variants, homologues and functional fragments thereof. Withdrawal of the rejection of claims 1-4, 7-17, 25, 28-44, 46-47, 49-50, 52-53, 79-81, 86-87, 90-92, 95-101 and 103-121 under 35 U.S.C. §112, first paragraph, is therefore warranted.

Claims 1-4, 7-17, 25, 28-44, 46-47, 49-50, 52-53, 79-81, 86-87, 90-92, 95-101 and 103-121 have been rejected under 35 U.S.C. § 112, first paragraph, as allegedly directed to subject matter which is non-enabled by the specification. It is the Examiner's position that the specification *does* enable methods for stimulating root growth, enhancing lateral or adventitious formation, effecting the expression of a polypeptide encoded by SEQ ID NO:26 and increasing the size of the root meristem, comprising transforming a plant with a nucleic acid sequence comprising SEQ ID NO:26 encoding SEQ ID NO:4, and vector, host cells, plant cells and plants transformed therewith. Office Action, page 8.

It is also the Examiner's position however, that the present invention allegedly does not reasonably provide enablement for methods of stimulating root growth or enhancing lateral root formation comprising transforming a plant with a nucleic acid hybridizing to SEQ ID NO:26, nucleic acids that are divergent from a nucleic acid

encoding SEQ ID NO:4, nucleic acids divergent from those nucleic acids as specified in claim 2 due to differences between alleles, functional fragments of said nucleic acids having biological activity of a cytokinin oxidase, a nucleic acid encoding an amino acid sequence comprising SEQ ID NO:32 and which is at least 70% similar to the amino acid sequence of SEQ ID NO:4, a nucleic acid encoding any immunologically active fragment of a cytokinin oxidase encoded by SEQ ID NO:26 or any immunologically active fragment encoded by any of the nucleic acids specified in claim 3, a nucleic acid encoding any functional fragment of any cytokinin oxidase encoded by SEQ ID NO:26 or any previously mentioned sequence. Office Action, page 8.

Enablement is a legal determination of whether a patent enables one skilled in the art to make and use the claimed invention, *Raytheon Co. v. Roper Corp.* 724 F.2d 951, 960, 220 USPQ 592, 599 (Fed. Cir. 1983), and is not precluded even if some experimentation is necessary, although the amount of experimentation needed must not be unduly extensive. *Atlas Powder Co. v. E.I. DuPont De Nemours & Co.*, 750 F.2d 1569, 1576, 224 USPQ 409, 413 (Fed. Cir.1984); *W.L. Gore and Associates v. Garlock, Inc.* 721 F.2d 1540, 1556, 220 USPQ 303, 315 (Fed. Cir. 1983). Nothing more than objective enablement is required and therefore it is irrelevant whether this teaching is provided through broad terminology or illustrative examples. *In re Marzocchi*, 439 F.2d 220, 223, 169 USPQ 367, 369 (CCPA 1971).

An analysis of whether the claims are supported by an enabling disclosure requires a determination of whether that disclosure contained sufficient information regarding the subject matter of the claims as to enable one skilled in the art to make and use the claimed invention. In order to establish a *prima facie* case on lack of enablement,

the examiner has the initial burden to establish a reasonable basis to question the enablement provided for the claimed invention. *See In re Wright*, 999 F.2d 1557, 1561-62, 27 USPQ2d 1510, 1513 (Fed. Cir. 1993) (Examiner must provide a reasonable explanation as to why the scope of protection provided by a claim is not adequately enabled by the disclosure). *See also In re Morehouse*, 545 F.2d 162, 192 USPQ 29 (CCPA 1976).

The Patent and Trademark Office Board of Appeals has stated:

The test [for enablement] is not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed to enable the determination of how to practice a desired embodiment of the invention claimed.

*Ex parte Jackson*, 217 USPQ 804, 807 (1982).

The threshold step in any enablement analysis is to determine whether the Examiner has met his burden of proof by advancing acceptable reasoning inconsistent with enablement. Factors to be considered by the Examiner in determining whether a disclosure would require undue experimentation have been summarized by the board in *Ex parte Forman*, 230 USPQ546,547 (Bd Pat App In 1986). These factors include: (1) the quantity of experimentation necessary, (2) the amount of direction or guidance presented, (3) the presence or absence of working examples, (4) the nature of the invention, (5) the state of the prior art, (6) the relative skill of those in the art, (7) the predictability or unpredictability of the art, and (8) the breadth of the claims. *In re Wands*, 858 F.2d 731, 737, 8 USPQ2d 1400, 1404 (Fed. Cir. 1988).

Applicants respectfully submit that the specification specifically describes the chemical structures of different nucleotide sequences encoding a cytokinin oxidase comprising the amino acid sequence as set forth in SEQ ID NO:4. Further, the specification indicates that well known methods may be used to screen for cytokinin oxidase activity. *See* specification e.g., at page 14, last paragraph, to page 15, line 7, page 93, lines 23-26: where the standard assay for cytokinin oxidase activity based on conversion of [2-3H] iP to adenine as described in Motyka et al. 1996 *Plant Physiol.* 112: 1035-1043 and Example 4, Table 7.

Figure 1 of the present application illustrates a comparison of gene structure of the *Arabidopsis AtCKX* genes 1 to 4 and the maize *CKX* gene. Page 92, lines 25-36 of the specification teach: “[t]he predicted proteins encoded by the *Arabidopsis AtCKX* genes show between 32% and 41% sequence similarity with the maize protein, while they show between 35% and 66% sequence similarity to each other. Because of this reduced sequence conservation, it is not clear *a priori* whether the *Arabidopsis AtCKX* genes encode proteins with cytokinin oxidase activity. An alignment of the *Arabidopsis AtCKX* predicted proteins 1 to 4 and the maize *CKX* gene is shown in Figure 2.”

The specification further teaches however, that using the standard assay for cytokinin oxidase activity (Motyka et al, 1996), transgenic plant lines expressing the different *Arabidopsis AtCKX* proteins indicate that *AtCKX1*, *AtCKX2*, *AtCKX3*, and *AtCKX4* all encode proteins having cytokinin oxidase activity. *See* specification, page 93, lines 22-28, and Table 6 (*AtCKX1 transgenic*); page 100, line 30, to page 101, line 3 and Table 7 (*AtCKX2 transgenic*); page 108, lines 20-24 and Table 8 (*AtCKX3 transgenic*); page 110, lines 10-14, and Table 9 (*AtCKX4 transgenic*).

Moreover, the specification teaches in a number of different places that the phenotypes observed for the different AtCKX transgenics is very similar. *See e.g.*, page 107, lines 18-24:

The phenotypes observed for *AtCKX2* transgenics were very similar but not identical to the *AtCKX1* transgenics, which were in turn very similar but not identical to the results obtained for the tobacco transgenics. This confirms the general nature of the consequences of a reduced cytokinin content in these two plant species and therefore, similar phenotypes can be expected in other plant species as well.

*See also*, specification, page 109, lines 4-9: “[t]he phenotypes generated by overexpression of the *ATCKX3* gene in tobacco and *Arabidopsis* were basically similar as those of AtCKX1 and AtCKX2 expressing plants, i.e., enhanced rooting and dwarfing” and page 111, lines 4-6: “[t]he phenotypes generated by overexpression of the AtCKX4 gene in tobacco and *Arabidopsis* were basically the same as those of AtCKX1 and AtCKX2 expressing plants, i.e., enhanced rooting, reduced apical dominance, dwarfing and yellowing of intercostals regions in older leaves of tobacco.”

What becomes immediately apparent therefore, is that the specification clearly enables methods *inter alia*, for stimulating root growth, enhancing lateral root formation, and effecting the expression of cytokin oxidase polypeptides using nucleotide sequences encoding cytokinin oxidases having *as little as between 35% and 66% amino acid sequence similarity to each other!* If the specification enables these methods using such divergent cytokin oxidase proteins, surely the same methods are enabled using nucleic acids hybridizing under medium stringency hybridization conditions to a nucleotide sequence as set forth in SEQ ID NO:26, nucleic acids which are divergent from a nucleic acid molecule encoding SEQ ID NO:4, nucleic acid molecules divergent from the nucleic

acid molecules recited in claim 2 due to differences in alleles, functional fragments of such nucleic acid molecules having cytokin oxidase activity, a nucleic acid molecule encoding an amino acid sequence comprising SEQ ID NO:32 and which is at least 70% similar to the amino acid sequence of SEQ ID NO:4, a nucleic acid molecule encoding an immunologically active fragment of a cytokinin oxidase encoded by SEQ ID NO:26 or any immunologically active fragment encoded by any of the nucleic acids specified in claim 3, a nucleic acid molecule encoding a functional fragment of a cytokinin oxidase encoded by SEQ ID NO:26 or any previously mentioned sequence.

The Examiner has relied on Kaminek et al. (1990) *Plant Physiol.* 93:1530-1538 for the notion that the state-of-the art teaches that not all cytokin oxidases are the same. See Office Action, page 11. In particular, the Examiner posits that Kaminek et al. teach isolated cytokin oxidases from different *Phaseolus* species as having different enzyme activities. Hare et al. (1994) *Physiologia Plantarum* 91:128-136 is cited for its teaching that substrate specificity varies; cytokinin oxidase from the moss *Funaria* has a high affinity for the cytokinin kinetin, whereas most plant cytokinin oxidases do not have a high affinity for kinetin. The same research article has been cited for teaching that cytokinin oxidases from *Dictyostelium discoideum* and *Saccharomyces cerevisiae* have a broader substrate specificity than most plant cytokinin oxidases.

Applicants respectfully submit that the teachings of Kaminek et al. and Hare et al. are irrelevant to the presently claimed invention. Applicants claims recite a plant cytokinin oxidase with particular chemical structures and the use thereof in yielding plants with altered phenotypes. The facts that cytokinin oxidases from different *Phaseolus* species have different enzyme activities and that plant cytokinin oxidases

might not have a high affinity for kinetin--are not in any way relevant to the presently disclosed and claimed invention. The present application teaches that AtCKX1, AtCKX2, AtCKX3, and AtCKX4 all possess cytokinin oxidase activity.

On page 11 of the Office Action, the Examiner posits: [t]he state-of -the art is such that one of skill in the art cannot predict which nucleic acids that exhibit less than 100% sequence identity to SEQ ID NO:26 will encode a protein with the same activity as a protein encoded by SEQ ID NO:26. It is respectfully submitted that the claims recite nucleic acid molecules encoding a plant cytokinin oxidase with particular chemical structures. The claims do not require, nor is there any basis in law for requiring that the claimed cytokinin oxidases have the same activity as the protein encoded by SEQ ID NO:26. What is required is that the presently claimed nucleotide sequences encode a protein having cytokinin oxidase activity and that the molecules encoding such a protein work in the presently claimed methods of e.g., stimulating root growth, enhancing the formation of lateral or adventitious roots, increasing seed size or weight, embryo size or weight, and cotyledon size or weight, as presently recited in the claims.

As fully discussed above, the present application clearly enables methods *inter alia*, for stimulating root growth, enhancing lateral root formation, and effecting the expression of cytokinin oxidase polypeptides using nucleotide sequences encoding cytokinin oxidases having *as little as between 35% and 66% amino acid sequence similarity to each other*. It certainly is predictable therefore, that nucleic acid molecules exhibiting less than 100% sequence identity to SEQ ID NO:26, including the presently claimed nucleic acids hybridizing under medium stringency hybridization conditions to a nucleotide sequence as set forth in SEQ ID NO:26, nucleic acids which are divergent



from a nucleic acid molecule encoding SEQ ID NO:4, nucleic acid molecules divergent from the nucleic acid molecules recited in claim 2 due to differences in alleles, functional fragments of such nucleic acid molecules having cytokin oxidase activity, a nucleic acid molecule encoding an amino acid sequence comprising SEQ ID NO:32 and which is at least 70% similar to the amino acid sequence of SEQ ID NO:4, a nucleic acid molecule encoding an immunologically active fragment of a cytokinin oxidase encoded by SEQ ID NO:26 or any immunologically active fragment encoded by any of the nucleic acids specified in claim 3, a nucleic acid molecule encoding a functional fragment of a cytokinin oxidase encoded by SEQ ID NO:26 or any previously mentioned sequence, may encode a cytokinin oxidase for use in the presently claimed invention.

The Examiner has also cited Bowie et al. (1990) *Science* 247:1306-1310 for teaching that prediction of protein structure from sequence data and, in turn, utilizing predicted structural determinations to ascertain functional aspects of the protein, is extremely complex. Bowie et al (1990) is also cited for teaching that positions within a protein's sequence where amino acid substitutions can be made with a reasonable expectation of maintaining function are limited. This same article however, also states that "many different sequences can code for proteins with essentially the same structure and activity," (page 1306, left column) and that "[s]tudies in which these methods [for studying protein tolerance to sequence variation] were used have revealed that proteins are surprisingly tolerant of amino acid substitutions" (page 1036, right column).

Thus, while perhaps complex, such a determination is not undue, but rather routine today as opposed to 1990 when Bowie et al. was published. Furthermore, "the specification in question provides a reasonable amount of guidance with respect to the

direction in which the experimentation should proceed” *In re Wands*, 858 F.2d at 737, a significant factor in demonstrating that experimentation is not undue..

Applicants respectfully submit that the Examiner has not considered for example, Figure 2 of the present application which clearly indicates regions of amino acid identities and similarities between maize (ZmCKX1) and *Arabidopsis* (AtCKX1 to AtCKX4) cytokinin oxidases. Thus, besides the teaching of the specification where different *Arabidopsis thaliana* cytokinin oxidases having as little as between 35% and 66% amino acid sequence similarity to each other function to impart the same phenotypic characteristics to transgenic plants, one skilled in the art would have ample guidance as to the preferred location of any amino acid substitutions.

Still further, pages 36-37 and Table 1 of the specification provide guidance in substituting amino acids having similar properties such as hydrophobicity, hydrophilicity, hydrophobic moment, antigenicity, propensity to form or break  $\alpha$ -helical structures or  $\beta$ -sheet structures. A well defined, standard cytokin oxidase activity assay is also provided (Motyka et al. 1996 *Plant Physiology* 112: 1035-1043) for testing transgenic plants expressing the claimed cytokinin oxidase proteins.

Citing McConnell et al. (2001) *Nature* 411 (6838):709-713, the Examiner states that replacement of a glycine residue located within the START domain of either the PHABULOSA or PHAVOLUTA protein receptor with either an alanine or aspartic acid residue, alters the sterol/lipid binding domain. The Examiner has not, however, explained how or why potential amino acid substitution within the START domain of either the PHABULOSA or PHAVOLUTA protein is related to or affects plant cytokinin oxidase expression.

Citing Fourgox-Nicol et al. (1999) *Plant Molecular Biology* 40:857-872, the Examiner gives an example of a hybridization experiment where isolation of a 674 bp fragment using a 497 bp probe under stringent hybridization conditions yielded a number of sequence differences such as a 99 bp insertion, single nucleotide gaps, nucleotide mismatches. Applicants again fail to see the relevance of this article to the presently claimed invention. One skilled in the art, having the present application in hand could easily perform a hybridization experiment and determine if a nucleic acid molecule hybridized to a nucleic acid molecule having the nucleotide sequence as set forth in SEQ ID NO:26 under various hybridization conditions, including medium stringency hybridization conditions as presently recited in the claims. Further, one skilled in the art could transform a plant with such a nucleotide sequence and assay for cytokinin oxidase activity using a standard assay as taught repeatedly in the specification. Moreover, having the teaching of the specification in hand that amino acid sequences having only between 35% and 66% amino acid sequence similarity to each other function as cytokinin oxidase proteins and impart the same phenotypic characteristics to transgenic plants, one skilled in the art would reasonably believe that sequences hybridizing to SEQ ID NO:26 under medium stringent conditions would encode cytokinin oxidase proteins and impart the same phenotypic characteristics to transgenic plants.

On page 8 of the Office Action, the Examiner has stated: "In addition, Applicants' disclosure does not reasonably provide enablement for methods for altering root geotropism, increasing the size of seeds, increasing embryo size, or increasing cotyledon size, increasing yield, altering leaf senescence, increasing leaf thickness, reducing vessel size, improving standability of seedlings, increasing branching, improving lodging

resistance, increasing early vigor and stress tolerance comprising transforming a plant with SEQ ID NO:26 encoding SEQ ID NO:4 or any nucleic acid encoding a fragment of SEQID NO:4 or a immunologically active fragment, or nucleic acid encoding a divergent polypeptide or a nucleic acid that hybridizes with SEQ ID NO:26, or any of the nucleic acids listed in claim 2 or 3 and Applicant is not enabled for a diagnostic composition.”

Applicants respectfully submit that with respect to increasing the size of seeds, increasing embryo size, or increasing cotyledon size, the specification is clearly enabling. For example, page 14, lines 8-12 of the specification teach:

In accordance with the present invention, it has also been surprisingly discovered that transgenic plants overexpressing a cytokinin oxidase gene develop seeds (including embryos) and cotyledons of increased size and/or weight. These results are surprising as a reduced cytokinin content would have been expected to be associated with a reduced organ growth.

Still further, Applicants direct the Examiner to Example 16, Table 11, and Figures 13 and 14 where increased embryo, cotyledon and seed size are specifically exemplified for transgenic plants expressing AtCKX1, AtCKX2, AtCKX3, and AtCKX4.

With respect to a diagnostic composition, claims 79-81 have been amended so that they no longer recite “diagnostic.”

On page 10 of the Office Action, the Examiner has stated: “[r]e Claims 1-2, 46-47, 90-92, 95-97, 103-108. applicants fail to provide guidance for methods of altering root geotropism, increasing yield, altering leaf senescence, increasing leaf thickness, reducing vessel size, improving standability of seedlings, increasing branching, improving lodging resistance, increasing early vigor, stress tolerance, seed size, embryo size or cotyledon size comprising transforming a plant with SEQ ID NO:26 encoding

SEQ ID NO:24, or transforming a plant with any nucleic acid that encodes a protein that lowers the level of cytokinins in a plant. In fact, Applicants disclose that plants transformed with SEQ ID NO:26 had smaller leaves (page 106, bottom table) and seeds whose weight was not statistically different from untransformed seeds (page 105, bottom table). Applicants have not disclosed how decreasing cytokinin levels alters root geotropism, increases leaf thickness or how yield is increased even though plants produce less seeds, smaller leaves and flower size is not changed. “

Specifically in response to these statements of the Examiner, Applicants respectfully submit the following. Although page 104, lines 21-22 of the specification, discloses “the weight of the individual seeds was comparable to the weight of seeds from wild type plants,” the very next page, lines 5-7 indicates: “[p]lease note that these plants were grown under green house conditions during winter time. This affects negatively the number of flowers that are formed, in particular in the transgenic clones. Further, as previously discussed hereinabove, Example 16 provides ample data evidencing increases in seed/embryo/cotyledon size. Additional data which further supports the teaching of the application in this regard is submitted herewith at Exhibit E.

Applicants further direct the Examiner to specific locations in the specification where other claimed phenotypes are thoroughly discussed. For example, one skilled in the art would understand root geotropism to mean the downward growth of roots, which results in longer primary roots and/or lateral or secondary roots. *See* page 15, lines 20-24; page 17, lines 10-17; page 18, lines 1-6; page 20, lines 4-8; page 28, lines 1-4; page 29, lines 12-16; page 32, lines 16-18; Example 3: page 96, lines 6-8; page 98, line 20 to page 99, line 13; page 100, lines 4-5; Example 4: page 102, lines 18-20; page 106, lines

1-15; p.107, lines 12-13; Example 5: page 109, lines 4-6; Example 6: p.111, lines 4-7; Example 9; page 113, lines 16-18.

Increased yield can be exemplified in various forms such as increased biomass of one or more organs such as increased root biomass, increased shoot biomass, increased seed biomass, but also as increased production of specific compounds. *See*: p.17, lines 10-17 and 23-28; page.25, lines 12-17; page 26, lines 21-25; page 29, lines 8-9; page 30, lines 4-7; page 32, lines 7-9; page 33, lines 13-20.

For specific disclosure enabling increased root growth, *see* Example 3: page 96, lines 1-16; page 98, line 20 to page 99, line 13; Example 4: page 102, line 15 to p.104, line 15; page 106, lines 1-23; Example 5: page 109, lines 4-6; Example 6: page 111, lines 4-7.

Increased shoot growth is specifically taught in Example 12, page .117, lines 11-17 (by grafting, but can also be obtained by using tissue specific promoters: *see* page 117, lines 28-29).

Altering leaf senescence is also amply supported by the teachings of the present application. *See* specification page 13, lines 22-26; page 28, lines 17-21; page 30, lines 4-7; page 33, lines 20-22. *See also* Figure 12A (CKX2 transgenic graft shows less senescence than wild type on the left); Example 3 page 95 lines 11-16 and Figure 7E (these are results for CKX1: the principle of using CKX for delaying senescence works also for homologues of CKX2; p.101, lines 15-19 (senescence well known in the art to be manifested by yellowing); page 107, lines 18-23, page 111, lines 4-6 and 13-15; Example 12: page 117, lines 11-13 and lines 25-27; Example 13: page 118, lines 12-13; Example 14: guidance for choosing a promoter.

Increased leaf thickness is specifically taught on page 28, lines 22-25; Example 10, page 114, lines 13-15 (increased size of leaf parenchyma and epidermal cells) and Figures 10C & D (data for CKX1)

Reducing vessel size is taught on page 28, lines 26-30; Example 10, page 114 lines 12-13 (size of phloem and xylem is reduced), *see also* Figure 10B (data for CKX1).

Improving standability of seedlings is taught on page 29, lines 4-7; page 82, lines 23-27, where a definition of standability is related to increased root growth. Increased root growth is demonstrated in Example 3, page 96, lines 1-16; page 98, lines 19-13; Example 4, page 102, line 13 to page 104, line 15; page 106, lines 1-22, Example 5, page 109, lines 4-9 (similar phenotype as in previous examples); Example 6: p.111, lines 4-8; Example 9: transgenic seedlings have increased root growth compared to wild type.

Increased branching is taught on page 22, lines 12-16; page 31, lines 25-27; Example 3, page 96, lines 6-8 (root branching), page 96, lines 26-30 and Figure 9 (shoot branching), page 98, lines 11-12 and page 99, lines 5-9 (root branching); Example 4, page 101, lines 15-19 (shoot branching), page 102, lines 18-20 (root branching), page 106, lines 15-20 (increased root branching), Example 9, page 113, items E and F (increased root branching).

Improving lodging resistance is taught on page 82, lines 23-27 (definition of lodging resistance: related to increased root growth). This same data also teaches improving standability.

Increasing early vigor is taught on page 17, lines 10-17; page 25, lines 12-19; page 26, lines 28 to page 27 line 4; p.28, lines 5-9; page 32, lines 23-27; page 82, lines

19-22, Examples 9 and 16 (increased seed size/weight, increased embryo size/weight, increased cotyledon size/weight).

Increasing stress tolerance is taught on page 25, lines 18-21; page 27, lines 1-2; page 30, lines 8-10; page 53, lines 25-26; page 64, table; p.70 lines 25-27.

With regard to the specific teachings outlined above, the specification also clearly teaches e.g., on page 107, lines 18-23; page 111, lines 10-20; and page 25, lines 22-26, that effects are not restricted to a particular type of CKX but are general for all CKX proteins.

Thus, upon reading the disclosure, one of ordinary skill in the art would have been provided with a reasonable amount of guidance to make and use nucleotide sequences encoding plant cytokinin oxidase proteins as recited in the presently amended claims for use in producing plants having altered phenotypes as also recited in the presently amended claims. The Examiner has not provided any evidence why one of ordinary skill in the art would not have had a reasonable expectation of success in practicing the presently claimed invention without having to engage in undue experimentation.

In view of the foregoing remarks, withdrawal of the rejection of claims 1-4, 7-17, 25, 28-44, 46-47, 49-50, 52-53, 79-81, 86-87, 90-92, 95-101 and 103-121 have been under the enablement provision of 35 U.S.C. §112, first paragraph, is warranted.

Claims 37 to 42 have been rejected under 35 U.S.C. §101 as allegedly directed to non-statutory subject matter. As suggested by the Examiner, claims 37 and 38 have been amended to recite in relevant part: “wherein the harvestable part comprises the nucleic acid molecule which was introduced into the transgenic plant.” Claims 41 and 42 recite in relevant part: “wherein the progeny comprises the nucleic acid molecule which was



introduced into the transgenic plant.” Withdrawal of the rejection of claims 37 to 42 under 35 U.S.C. §101 is therefore respectfully requested.

Claims 1-4, 7-17, 25, 28-44, 49-50, 52-53, 87 and 98-101 have been rejected under 35 U.S.C. § 102(b) as allegedly anticipated by Morris (February 1999 WO 99/06571. As presently amended, claims 2 and 3 have been amended to recite hybridization conditions associated with specific hybridization, e.g., medium stringency conditions. As presently amended, claims 2 and 3 recite in relevant part: “under medium stringency conditions such as 1-4X SSC/0.25 % w/v SDS at 45° C or higher for 2-3 hours.” Since the amino acid identity between the cytokin oxidase of the present invention having the amino acid sequence as set forth SEQ ID NO:4 encoded by SEQ ID NO:26 and the of cytokinin oxidase of Morris et al. is only 38.4% (*see* specification, page 88, lines 14-16), a nucleotide sequence encoding the cytokinin oxidase of Morris et al. would not hybridize to the nucleotide sequence set forth in SEQ ID NO:26 under the medium stringency conditions presently recited in the claims.

For example, AtCKX2 has a GC content of 43%. If the melting temperature in 5X SSC (which is 0.825 M Na<sup>+</sup>) without formamide is calculated according to the formula on page 50 of the specification:

$$T_m = 79.8^{\circ}\text{C} + (18.5 \times \text{Log}[\text{Na}^+]) + (58.4^{\circ}\text{C} \times \%[\text{G}+\text{C}]) - (820 / \# \text{ bp in duplex}) - (0.5 \times \% \text{formamide}),$$

then we have for a duplex of 100 bp, a melting temperature of 95.2°C and for a duplex of 500 bp, a T<sub>m</sub> of 101.7°C.

It is known in the art that for each % base mismatch, the melting temperature will drop with 1°C. *See* article by R. Deaton (provided herewith at Exhibit F), taken from R. Deaton’s website (<http://csce.uark.edu/~rdeaton/>), published in Proceedings of the Genetic

and Evolutionary Computation Conference, Volume 2, -, AAAI, Morgan Kaufmann, San Francisco, 1999.Orlando, FL, July 1999), page 2, 1<sup>st</sup> column, lines 13-15 below formula (2), and the “Nonradioactive In Situ Hybridization Application Manual from Roche , page 35 (provided herewith as Exhibit G).

The maize sequence has only 38.4% sequence identity with AtCKX2, which is 61.6% difference, hence the hybridization temperature will be 61.6°C lower: for a 100 bp duplex: 33.6°C; for a 500 bp duplex: 40.1°C.

Therefore, under medium stringency conditions (4x SSC and 45°C or higher), no hybridization will occur. Hybridization under medium stringency conditions as presently recited in the claims, would therefore exclude the maize CKX taught by Morris (February 1999 WO 99/06571). Based on the foregoing, the presently claimed invention is therefore distinguished from Morris et al.

With respect to claim 1, it is submitted that Morris does not teach a method for producing a plant having enhanced root growth or for producing a plant having enhanced formation of primary, lateral, or adventitious roots as presently recited in the claim.


It is axiomatic that anticipation under section 102 requires that the prior art reference disclose *every element* of the claims. *In re King*, 801 F.2d 1324,1326, 231 USPQ 136, 138 (Fed. Cir. 1986). Thus, there must be *no* difference between the subject matter of the claim and the disclosure of the prior art reference. Stated otherwise, the reference must contain within its four corners adequate direction to practice the invention. The corollary of this rule is equally applicable. The absence from the

reference of *any* claimed element negates anticipation. *Kolster Speedsteel AB v. Crucible, Inc.*, 793 F.2d 1565, 1571, 230 USPQ 81, 84 (Fed. Cir. 1986).

As supported by the holding of *Kolster Speedsteel*, Morris falls far short of the statutory standard of 35 U.S.C. §102(b). Withdrawal of the rejection of claims 1-4, 7-17, 25, 28-44, 49-50, 52-53, 87 and 98-101 under 35 U.S.C. § 102(b) is therefore respectfully requested.

Accordingly, in view of the amendments to the claims, the abstract, the Sequence Listing and the foregoing remarks and attached exhibits, the present application is believed to be in condition for allowance, which action is earnestly solicited.

Respectfully submitted,

A handwritten signature in black ink, appearing to read "Ann R. Pokalsky". The signature is fluid and cursive, with the first name "Ann" and last name "Pokalsky" clearly distinguishable.

Ann R. Pokalsky  
Reg. No. 34,697  
Attorney for Applicants

DILWORTH & BARRESE  
333 Earle Ovington Blvd.  
Uniondale, NY 11553  
(516) 228-8484  
ARP:bg

## Cytokinin oxidase from *Zea mays*: purification, cDNA cloning and expression in moss protoplasts

Nicole Houba-Hérin<sup>1,†</sup>, Claude Pethe<sup>1,†</sup>, Jacques d'Alayer<sup>2</sup> and Michel Laloue<sup>1,\*</sup>

<sup>1</sup>Laboratoire de Biologie Cellulaire INRA, Route de St-Cyr, 78026 Versailles Cedex, France, and

<sup>2</sup>Laboratoire de Microséquençage des Protéines, Institut Pasteur, 25, rue du Dr Roux, 75724 Paris Cedex 15, France

### Summary

Cytokinins are degraded by cytokinin oxidases (CKOs) which catalyse cleavage of the *N*<sup>6</sup>-(isopent-2-enyl)-side chain resulting in formation of adenine-type compounds. CKO activity has been recorded in many plants and is thought to play a key role in controlling cytokinin levels in plants. Several partially purified CKOs have been characterised but no genes have been isolated yet. CKO activity is known to be inhibited by phenylureas, cytokinin agonists. We used 1-(2-azido-6-chloropyrid-4-yl)-3-(4-[<sup>3</sup>H])phenylurea ([<sup>3</sup>H]-azidoCPPU) to photolabel a glycosylated CKO from maize kernels. This enabled us to purify the enzyme. Peptide sequences were determined and the corresponding cDNA was cloned. The deduced amino acid sequence shares homology domains with FAD-dependent oxidases. An original assay based on transient expression of the enzyme in moss protoplasts allowed the functionality of the recombinant enzyme to be demonstrated.

### Introduction

Cytokinins are hormones that control plant growth and development including growth of lateral buds, leaf expansion and delay of leaf senescence. Together with auxins, cytokinins are required for cell division and shoot formation in tissue culture (Miller *et al.*, 1956; Skoog and Miller, 1957). Naturally occurring cytokinins are *N*<sup>6</sup>-substituted adenine derivatives and their metabolism in plants is well established (McGaw and Burch, 1995; for a review). However, the initial step(s) of their biosynthesis in plants are still unknown. A gene coding for the enzyme that adds an isoprenyl chain on the amino group of AMP leading to a cytokinin molecule has only been found so far in some phytopathogenic bacteria, like *Agrobacterium* and *Pseudomonas* (Morris, 1995).

In tobacco callus and cell suspensions, the active cytokinins are thought to be the free bases (Hecht *et al.*, 1975 and Laloue, 1980, respectively). The base, riboside and nucleotide forms of cytokinins can be interconverted by purine metabolism enzymes (Chen, 1981). Some of the corresponding genes recently have been cloned and sequenced from *Arabidopsis thaliana* and the moss *Physcomitrella patens* (Moffatt *et al.*, 1992; Schnorr *et al.*, 1996; von Schwartzberg *et al.*, 1998).

Irreversible or transient cytokinin inactivation can proceed through conjugate formation. *N*-glucosylation and alanine conjugation lead to biologically inactive and very stable cytokinins (Mok and Martin, 1994). Martin *et al.* (1997) recently cloned from *Phaseolus vulgaris* seeds the gene coding for the *O*-xylosyltransferase, an enzyme which forms *O*-xylosylzeatin from zeatin. The cytokinin-*O*-glycoside conjugates can be hydrolysed into more active cytokinins and are assumed to serve as storage products. Brzobohaty *et al.* (1993) demonstrated the release of free cytokinin from cytokinin-*O*-glucoside conjugates by a  $\beta$ -glucosidase encoded by a cloned *Zea mays* cDNA.

In addition, cytokinins can be irreversibly inactivated by cytokinin oxidase (CKO) which oxidatively cleaves the *N*<sup>6</sup>-side chain of cytokinin bases and ribosides to yield adenine and adenosine, respectively (Whitty and Hall, 1974). Cytokinin degradation indeed was observed to be a significant component of cytokinin metabolism in tissues of many plants supplied with radiolabelled cytokinins. In tobacco cells, exogenously supplied *N*<sup>6</sup>-( $\Delta^2$ -isopentenyl)adenosine (IPA) is readily catabolized with a half-life of 3 h (Terrine and Laloue, 1980). Therefore, changing the expression of CKO should be a very powerful tool for modulating cytokinin levels in plants. CKOs have been partially purified from three main plant sources: maize kernels (Burch and Horgan, 1989; Meilan and Morris, 1994) or seedlings (Burch and Horgan, 1992; Schreiber *et al.*, 1995), wheat germ (Laloue and Fox, 1989) and *Phaseolus* callus tissue (Chatfield and Armstrong, 1988). CKOs are generally described as glycoproteins. Highly diverse molecular weights were reported for CKOs in different plants (i.e. from 25.1 to 94 kDa) or even in the same plant (Armstrong, 1994; for a review). However, as yet, no genes have been isolated.

Some diphenylureas which are very potent cytokinin agonists have been shown to inhibit CKO activity (Laloue and Fox, 1985, 1989). This property has been largely confirmed and extended to other urea derivatives like thidiazuron (Chatfield and Armstrong, 1986). This inhibitory effect has been shown to be non-competitive (Burch and Horgan, 1989; Wang and Letham, 1995).

Received 14 October 1998; revised 12 January 1999; accepted 13 January 1999.

\*For correspondence (fax +33 130833099; e-mail laloue@versailles.inra.fr).

<sup>†</sup>These authors contributed equally to the work presented here.

Table 1. Purification of maize cytokinin oxidase

Steps	Activity (nmol min <sup>-1</sup> ) <sup>a</sup>	Volume (ml)	Proteins (mg)	Specific activity (pmol mg <sup>-1</sup> min <sup>-1</sup> )	Yield (%)	Purification fold
Crude extract	126.0	860	2236.0	56.3	100.0	–
30–50% ammonium sulfate fraction	56.0	65	1105.0	50.7	44.0	–
ConA-Sepharose chromatography	18.1	38	15.7	1153.0	14.0	20.5
ResourceQ FPLC	2.9	28	0.3	9206.0	2.3	163.5

<sup>a</sup>nmol of iPA degraded per min at 30°C (pH 7.5). Assays performed with [<sup>3</sup>H]-iPA 2 µM.

We developed an azido-derivative of *N*-(2-chloro-4-pyridyl)-*N'*-phenylurea (CPPU) (Dias *et al.*, 1995) and showed that this probe can be used to photolabel and purify cytokinin-binding proteins (Gonneau *et al.*, 1998; Nogué *et al.*, 1996). As expected, tritiated azidoCPPU was shown to photolabel CKO which was then purified by 2D-gel electrophoresis. We report here the cloning of a cDNA for maize CKO. Amino acid sequencing of internal peptides made it possible to design degenerate oligonucleotides. The cDNA was recovered and sequenced. The activity of the encoded protein was demonstrated using a transient expression assay in moss protoplasts.

## Results

### [<sup>3</sup>H]-azidoCPPU labelling of cytokinin oxidase (CKO)

**Partial purification of a maize CKO.** CKO enrichment starting from 400 g of maize kernels is presented in Table 1. The initial purification procedure was adapted from Chatfield and Armstrong (1988) and Burch and Horgan (1989). After anion exchange chromatography the yield in active CKO was 2.3% with a purification factor of at least 160. However, no discrete peak corresponding to the activity was obtained indicating that the enzyme was not purified to homogeneity. Analysis by SDS-PAGE also confirmed that the active fraction contained a protein mixture (data not shown).

**Photolabelling of maize CKO.** The partially purified post-conA Sepharose fraction of maize CKO was exposed to [<sup>3</sup>H]-azidoCPPU and irradiated for promoting covalent binding of the probe to affine proteins. SDS-PAGE analysis and fluorography then showed a radiolabelled band corresponding to a molecular weight of 63 kDa (data not shown). Competition experiments using substituted ureas like CPPU and purine-type cytokinins which are substrates of the enzyme led to a displacement of the photolabel (data not shown). The post-conA labelled fraction was analysed further by FPLC anion exchange and gel permeation chromatography, respectively (Figure 1a,b). In both cases, we noticed that the CKO activity peak was exactly superimposed on the radioactive peak. These cochromatography

results indicated that the labelled protein most likely corresponded to CKO.

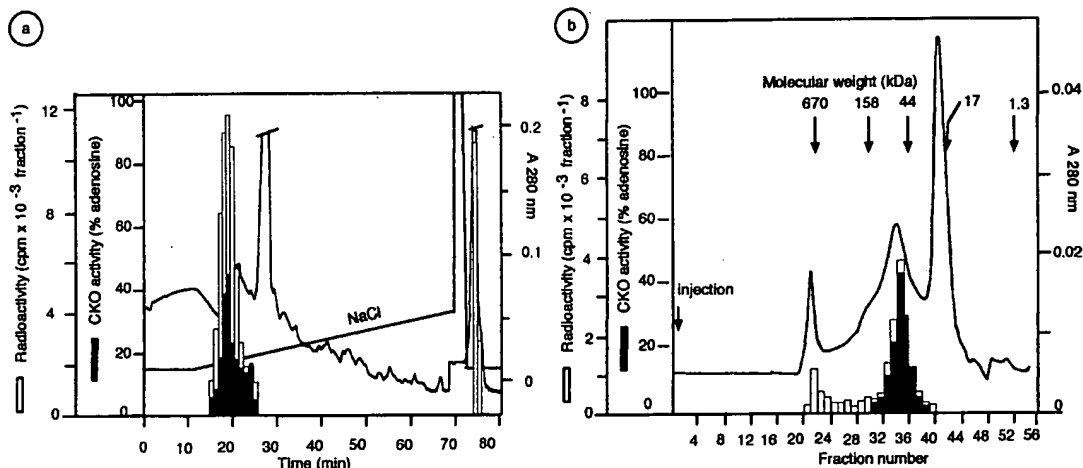
**Functional evidence that the labelled protein is CKO using a suicide substrate.** Suicide substrates are mechanism-based enzyme inactivators and acetylenic and/or allenic groups are classical enzymatically activable functional groups (for details, see Walsh, 1982). HA8, an *N*<sup>6</sup>-substituted adenine with a C<sub>4</sub> allenic chain (–CH<sub>2</sub>–CH = C = CH<sub>2</sub>), is the most active of a series of such suicide substrates which inhibit CKO activity in a time and concentration dependent fashion (C. Pethe, unpublished results). The design of these inhibitors was reasoned on the basis of the formation, in the active site, of an enamine in the isoprenic chain, as demonstrated with wheat CKO (Laloue and Fox, 1985). Thus oxidation of HA8 by the enzyme should lead to an alleneimine intermediate of greater nucleophilicity expected to allow the addition of nucleophilic groups of the enzymatic site.

As illustrated in Figure 2, inactivation (87% after 2 h incubation) of maize CKO in the presence of HA8 resulted in a 77% reduction of the photolabelling (lane 2) as compared to control samples which had been incubated either without any substrate (lanes 1 and 4) or with *N*<sup>6</sup>-(Δ<sup>2</sup>-isopentenyl)adenine (iP) (lane 3). Thus, inhibition of the labelling is strictly related to the inactivation of the enzyme since HA8, without preincubation and at the final concentration of 0.1 µM, had a limited effect on CKO photolabelling (lane 4), as previously established in competition experiments (results not shown). iP was without any significant effect (lane 3), because it was entirely degraded by CKO during the course of the incubation.

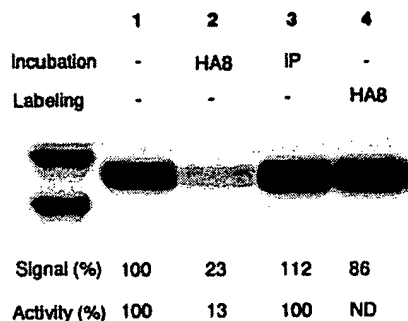
Hence, these results show that the labelling of the protein with the probe [<sup>3</sup>H]-azidoCPPU is significantly reduced when the enzyme has previously been inactivated through a mechanism-based reaction with a specific suicide substrate. It can be concluded that the labelled protein indeed is CKO. This probe was then used to trace CKO through the last steps of its purification.

### [<sup>3</sup>H]-azidoCPPU-mediated purification of maize CKO

Maize CKO photolabelled with [<sup>3</sup>H]-azidoCPPU after conA-Sepharose chromatography was then purified on a



**Figure 1.** CKO activity coincides with [<sup>3</sup>H]-azidoCPPU photolabelled proteins after FPLC anion exchange separation and gel permeation. An aliquot of post-conA maize CKO fraction (activity 0.7 nmol min<sup>-1</sup> ml<sup>-1</sup>) was photolabelled. (a) A 2-ml aliquot was applied to a ResourceQ column which was eluted with a NaCl gradient. Fractions of 2 ml were collected. (b) A 100-μl aliquot of the same fraction was applied to a TSKSW gel permeation column eluted at a flow rate of 1 ml min<sup>-1</sup> and fractions of 250 μl were collected. In both cases, fractions were assayed for CKO activity and radioactivity.



**Figure 2.** CKO photolabelling with [<sup>3</sup>H]-azidoCPPU is reduced after specific inactivation of the enzyme with the allelic suicide substrate HA8. Four aliquots (0.4 ml) of a post-conA fraction of maize CKO were incubated at 30°C for 2 h: samples 2 and 3, in the presence of HA8 and IP 0.4 μM, respectively; samples 1 and 4, without any substrate. Then all samples were diluted fourfold with buffer. Each sample was then photolabelled with [<sup>3</sup>H]-azidoCPPU 0.2 μM after 5 min of further incubation at 4°C in the presence of the probe alone (samples 1–2–3) and the probe plus HA8 0.1 μM (sample 4). All samples were then analysed by SDS-PAGE and fluorography. Cytokinin oxidase activity of each diluted sample was measured in parallel. Upper and lower signals on the left lane correspond to the 69 and 46 kDa [<sup>14</sup>C]-labelled molecular weights. ND = not done.

ResourceQ column as illustrated in Figure 1(a). The labelled fractions were further characterised by 2D-gel electrophoresis. Fluorography revealed the presence of one major labelled spot of approximately 63 kDa (Figure 3b). Silver staining showed a major protein spot and minor spots having the same molecular weight of about 63 kDa but differing in pI (Figure 3c). The labelled spot corresponded to one of the minor spots and was weakly stained by Coomassie blue (Figure 3a) suggesting that CKO was still a minor component of the extract.

At this stage, we considered the possibility that the

electrophoretic behaviour of the azidoCPPU-modified enzyme could differ from that of the unmodified enzyme. This was indeed the case, as shown after protein staining of the 2D-electrophoresis gel (Figure 3d) of the post-ResourceQ fraction treated with 5 μM cold azidoCPPU (3 rounds). The spot corresponding to the radiolabelled protein and indicated by an arrow on the figure had indeed increased in intensity (Figure 3d) as compared to the untreated sample (Figure 3c).

We exploited this modification in the electrophoretic behaviour of the azidoCPPU-treated CKO to purify the enzyme. A total of seven preparative 2D-gels were run and stained with amido-black. The spots of interest were recovered, pooled and submitted to amino acid sequencing after endoproteolytic cleavage since the N-terminus of the protein was blocked. The digested sample amounted to about 6 μg protein. Four internal peptide sequences were obtained (Figure 4): WNRFFVEMK (peptide 1); RLLSPGQDIFN (peptide 2); TYLARHTDRSDWVRHFGAAK (peptide 4) and GILQGTDIVGLIVYPLNK (peptide 5).

#### Isolation of a CKO cDNA

Poly A<sup>+</sup> RNA was prepared from maize cobs harvested about two weeks after anthesis. RT-PCR was performed by using degenerate oligonucleotides designed in both orientations from the four peptide sequences. Twelve oligonucleotide combinations were tested in 'touchdown' PCR to uncover all possible orders of the peptides. One combination resulted in an amplification product, i.e. a PCR reaction with primers designed from the sequences of peptides 5 and 1 (Co51 and Coa1) resulted in an

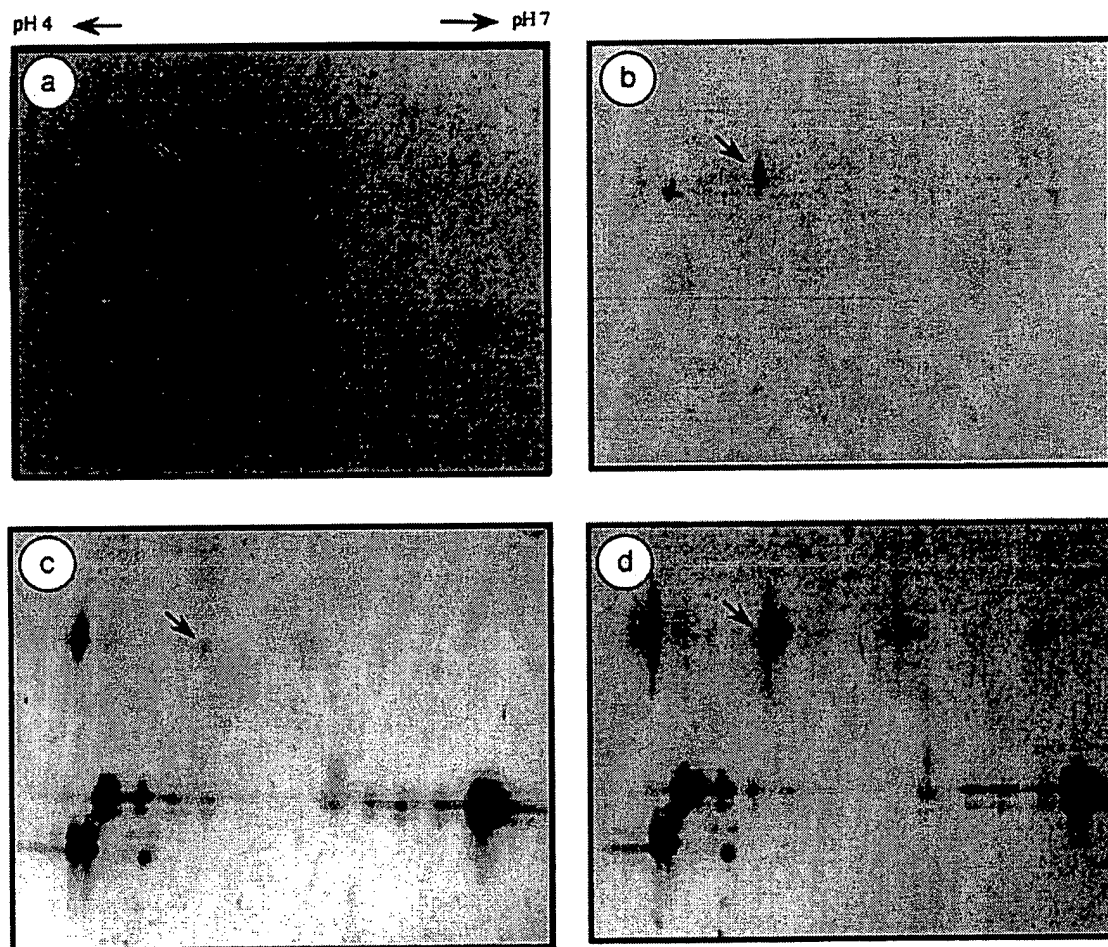


Figure 3. Localization of maize photolabelled CKO by 2-D gel electrophoresis and fluorography.

Half of an 8-ml aliquot of the post-conA fraction was photolabelled with [ $^3\text{H}$ ]-azidoCPPU 0.5  $\mu\text{M}$  and purified by anion exchange chromatography (as in Figure 1a) and analysed by 2D-gel electrophoresis. The equivalent of 3.7 ml of the post-conA fraction was applied to the gel stained with Coomassie blue (panel a) and analysed by fluorography (panel b). The equivalent of 0.3 ml was also analysed and the gel silver stained (panel c). The remaining half of the photolabelled post-conA fraction purified by anion exchange chromatography was further treated with cold azidoCPPU 5  $\mu\text{M}$  (3 rounds) and a 1/10th aliquot was analysed by 2D-gel electrophoresis and silver staining of the gel (panel d). Arrows indicate the position of the azidoCPPU-modified CKO.

approximately 300 bp amplified fragment by using as a template an aliquot of an initial PCR reaction performed with primers designed from the sequences of peptides 5 and 2 (Co51 and Coa23). This DNA fragment was cloned and four clones were sequenced. The 297 bp long sequences were all identical. The deduced amino acid sequence is bordered by the parts of peptides 5 and 1 which were used to design the oligonucleotides. This sequence also contains the C-terminus of peptide 5 and the whole of peptide 4 (Figure 4).

By using specific primers designed within this 297 bp long sequence, RACE PCR was performed on poly A<sup>+</sup> RNA so that both 3'- and 5'-RACE reaction products would overlap over a 81 bp stretch. A CKs7-AP2

(approximately 400 bp) 3'-RACE reaction product and a CKa7-AP1 (approximately 1600 bp) 5'-RACE reaction product were recovered (AP1 and 2 are the Clontech anchored adaptors for RACE PCR). They were cloned and sequenced. At least two different clones and the two complementary strands contributed to the sequence determination for each RACE reaction product. The reconstructed nucleotide sequence is given in Figure 4. Its linearity was confirmed by sequencing the independently amplified PCR product corresponding to the ORF region (see below). It has an open reading frame of 1602 bp that codes for a protein of 534 amino acids with a calculated molecular mass of 57.2 kDa which comprises all four peptide sequences. Since there is a stop codon upstream of the first ATG in all three

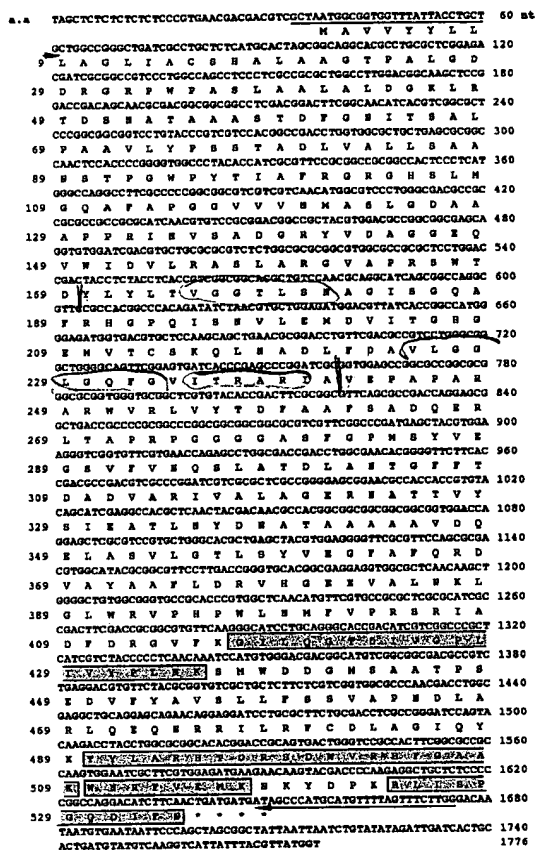


Figure 4. Nucleotide and deduced amino acid sequences of the CKO cDNA. The nucleotide sequence is presented in the top line and the deduced one letter amino acid sequence is shown below. The sequenced peptides are boxed. ORF-flanking oligonucleotides are indicated by arrows. The initial 297 bp sequence corresponds to nucleotides 1289–1585. Amino acid (a.a.) and nucleotide (nt) numbering is given at the left and the right side of the figure, respectively. (EMBL accession number is Y18377.)

reading frames, it is very likely that the ORF does not extend 5' to the longest 5'-RACE clone found.

The deduced CKO amino acid sequence shows sequence similarity with a FAD-binding domain (a.a. approximately 170–240) found in several oxidases, suggesting that CKO is a FAD-dependent oxidase (Figure 5). A GHS motif is also found in the CKO sequence (a.a. 104–106). This motif is typical for those enzymes that covalently bind to FAD through a histidine residue. A high sequence identity (30%) was observed with a hypothetical 47.9 kDa oxidoreductase encoded by the *FAS* operon of *Rhodococcus fascians* and the best match (42% identity) was with a hypothetical protein of *Arabidopsis thaliana* (AC002510).

#### Southern and Northern hybridisation analyses

A 0.6 kb long cDNA fragment (a 3' segment downstream of the *SacI* site at position 1082 in Figure 4) was used as

a probe for Southern and Northern hybridisations. Under low and high stringency washing conditions, the probe hybridises to one major fragment of genomic DNA for all the enzymes tested (Figure 6a,b). It was checked by restriction analysis of PCR-amplified genomic DNA that these enzymes do not cut the gene, *HindIII* excluded, at least in the region extending between the translation start and stop codons (data not shown). Some weakly hybridising bands were observed in all lanes even after high stringency washing indicating the presence of related genes in the maize genome. This was also observed when using an internal 1.1 kb long *NotI*-*BamHI* cDNA fragment (bp 340 to bp 1464 in Figure 4) (data not shown).

Total RNA and mRNA were purified from 2-week-old cobs. They were submitted to Northern analysis using the same probes. Figure 6(c) shows that the CKO mRNA is about 1.8 kb long, a size which is compatible with the longest reconstructed cDNA (1776 bp). The transcript is not abundant and is only clearly detected in a poly A<sup>+</sup> RNA population.

#### Transient expression of the cloned CKO in moss protoplasts

As an expression system that should be appropriate for the production of a correctly folded glycosylated plant protein, we used a transient expression assay in *Physcomitrella patens* protoplasts. Moss protoplasts were chosen because they do not require exogenous auxins and cytokinins for growth and division in contrast to higher plant protoplasts. The ORF-corresponding cDNA region was amplified from poly A<sup>+</sup> RNA with primers spanning the start and stop codons (CKs8 and CKa8) and cloned. One clone, whose sequence was identical to the sequence presented above, was subcloned into the expression vector pLBR19. Protoplast aliquots were transformed with pLBR19 carrying the CKO- or the  $\beta$ -glucuronidase-ORF (GUS-ORF), respectively. The GUS-ORF-transformed cells were used as a control for the CKO-ORF transformed cells and vice versa. Two days after transformation, CKO activity was measured and detected in the culture medium of the CKO-ORF transformed cells (Table 2). Only a weak activity was recovered in extracts made from these CKO-ORF transformed cells (data not shown). Incubation of [<sup>3</sup>H]-iPA in the presence of the culture medium from the control samples for up to 2 h did not lead to any adenosine production. GUS activity was assayed in the cells as a measure for the transformation efficiency of the protoplast batches. These results show that a catalytically active maize CKO is expressed in the transformed moss protoplasts.

#### Discussion

Cytokinin oxidase is a key and specific enzyme of cytokinin metabolism. Evidence presented in this paper indicates



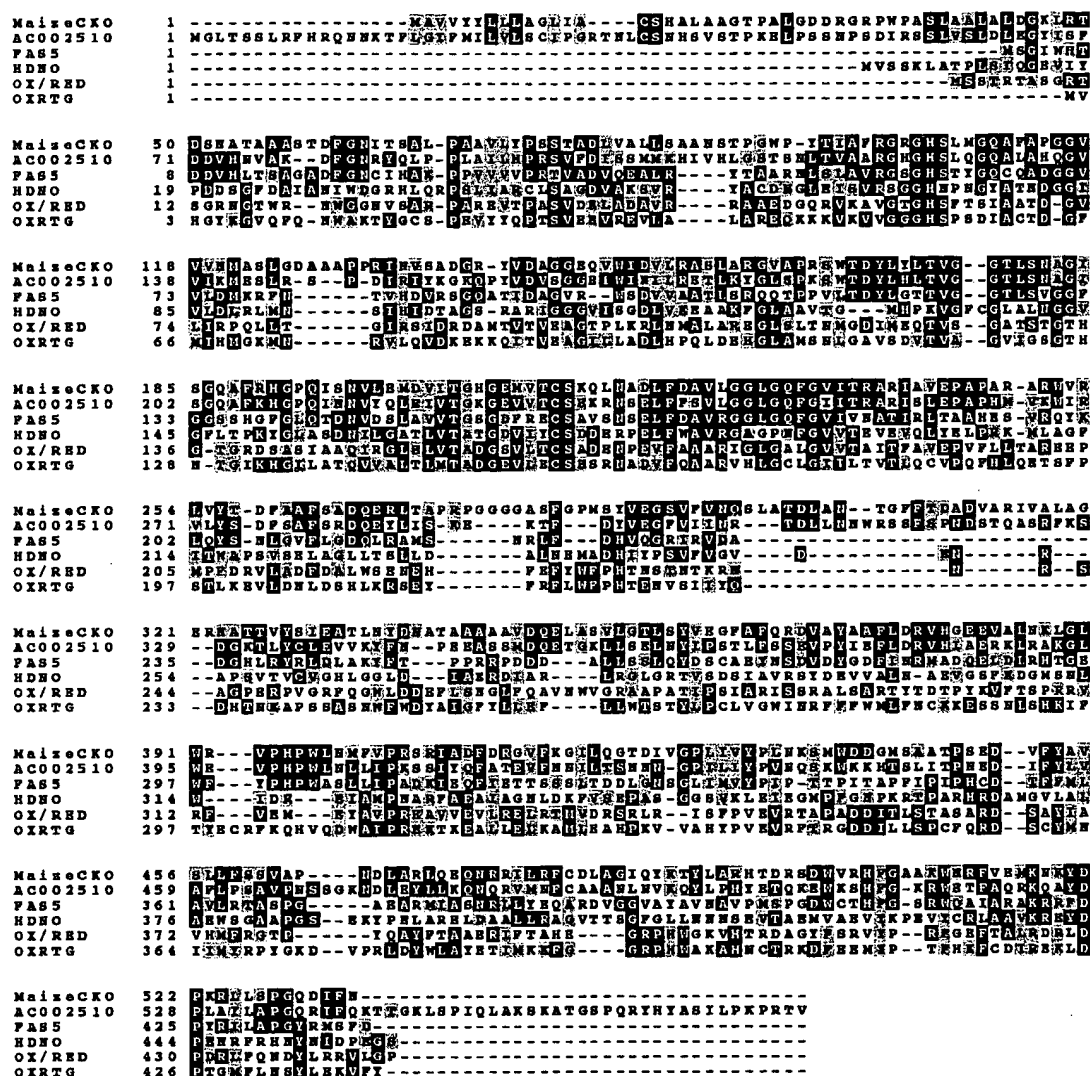


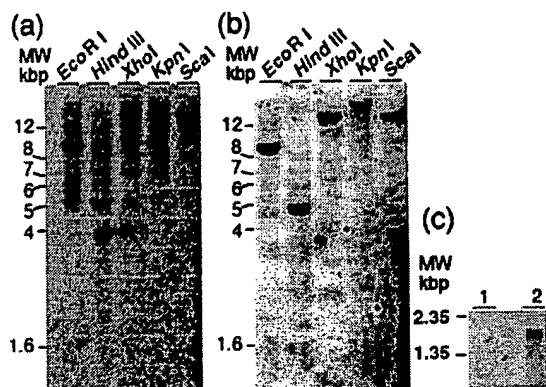
Figure 5. Alignment of the CKO-deduced amino acid sequence with several FAD-dependent oxidases.

Shaded regions indicate identical (black) or similar (grey) amino acids. AC002510: hypothetical protein in *Arabidopsis thaliana*; FAS 5: hypothetical 47.9 oxidoreductase in *Rhodococcus fascians* (P46377); HDNO: *Arthrobacter oxidans* 6-hydroxy-D-nicotine oxidase (X05999); OX/RED: FAD-dependent oxidoreductase in *Streptomyces coelicolor* (AL009204); OXRTG: *Rattus norvegicus* L-gulonolactone oxidase (P10867).

unambiguously that we have purified and cloned a maize cytokinin oxidase. Purification of the enzyme was assisted by the use of tritiated azidoCPPU which was shown to photolabel the active enzyme and to displace its electrophoretic migration towards a more basic pH. The probe-modified CKO could then be isolated from a set of proteins having the same molecular weight. Degenerate oligonucleotides related to the amino acid sequences of internal peptides allowed us to amplify a 297 bp cDNA fragment which was cloned and sequenced. A complete cDNA sequence could then be obtained after 5'- and 3'-RACE PCR using specific oligonucleotides. Finally, the cytokinin

oxidase ORF was cloned in a plant expression vector and its activity was clearly demonstrated in a transient expression assay using *Physcomitrella patens* protoplasts.

Photolabelling of the glycosylated CKO was very instrumental in identifying the protein after SDS-PAGE. Under such conditions, the protein has a molecular mass of 63 kDa which differs significantly from the previously reported 78 kDa for maize kernels CKO (Burch and Horgan, 1989) and 70 kDa for maize seedlings CKO (Schreiber *et al.*, 1995). These determinations are quite different from the 44 kDa molecular mass of the native enzyme measured by gel filtration using FPLC columns (Figure 1b; Burch and



**Figure 6.** Gene copy number and transcript size for the CKO gene. Maize var. Nobilis genomic DNA (8 µg) was digested with the indicated restriction endonucleases. The Southern blot was hybridized with a radiolabelled 0.6 kb probe (bp 1082–1675 in Figure 4). The membrane was washed either at low stringency (a) or at high stringency (b). RNA was purified from maize var. Nobilis cobs harvested 2 weeks after anthesis. Total RNA (10 µg, lane 1) and poly A<sup>+</sup> RNA (1 µg, lane 2) were resolved by gel electrophoresis and the Northern blot (c) was hybridised with the 0.6 kb probe. DNA and RNA length markers are given at the left of each panel in kbp.

**Table 2.** Transient expression of maize cytokinin oxidase in the culture medium of moss protoplasts

Protoplast batch <sup>a</sup>	CKO activity (%) <sup>b</sup>		GUS activity (AU min <sup>-1</sup> ) <sup>c</sup>	
	GUS-ORF	CKO-ORF	GUS-ORF	CKO-ORF
1	0	100	ND	0
2	0	100	3338	0
3	0	58	1075	0
4	0	75	1365	0

<sup>a</sup>Batches 1, 2 and 3 consisted of one sample per construct; batch 4 consisted of three samples per construct.

<sup>b</sup>Assays performed with [<sup>3</sup>H]-iPA 0.05 µM for 30 min at 30°C. Results are given as the percentage of iPA conversion into adenosine.

<sup>c</sup>GUS activity is expressed in arbitrary units per min using 4-methylumbelliferyl-β-D-glucuronide as a substrate. Results are given for 3.6 × 10<sup>5</sup> initially treated protoplasts.

AU = arbitrary unit; ND = not done.

Horgan, 1989). Glycosylated peptides often exhibit higher apparent sizes on SDS-PAGE gels because they bind less SDS molecules. Variations in the CKO glycosylation pattern could explain the discrepancy between the previously reported sizes of maize CKO observed on SDS-PAGE gels. In these studies, band identification as CKO was further supported by the observation that polyclonal antibodies raised against CKO purified fractions cross-reacted with CKO activity (Burch and Horgan, 1989; Schreiber *et al.*, 1995). However, it is also possible that monitoring the last steps of CKO purification by SDS-PAGE was misleading and that the observed interaction between the antibodies and CKO activity was due to the highly antigenic glycosidic

residues present on both CKO and the 70 or 78 kDa protein, respectively.

Preliminary attempts to express recombinant CKO in *Escherichia coli* cells were not successful and we therefore chose to analyse the functionality of the enzyme by using a plant system that should process the protein correctly. Moss protoplasts have the advantage of not requiring auxins and cytokinins for growth and division, although cytokinins are important in controlling moss development (Ashton *et al.*, 1979). Activity was mainly recovered in the culture medium. Either the enzyme is liberated in the medium by cell lysis or it is excreted. Therefore, functionality of the expressed recombinant enzyme is the final proof that we purified and cloned a maize CKO.

The reaction catalysed by CKO requires molecular oxygen and there is evidence that the reaction involves an iminopurine intermediate (Laloue and Fox, 1985). CKO is generally believed to be a copper-containing oxidase (Armstrong, 1994 for a review). However, no conclusive experimental data support this assumption. The deduced amino acid sequence of maize CKO cDNA has a domain (from position ≈ 170 to position ≈ 240) which is approximately 50% similar to a domain found in FAD-dependent oxidases. Furthermore, it has a conserved GHS motif (position 104/106) involved in the binding of FAD to a histidine residue in these flavoproteins (Figure 5). The highest homology is observed with the translated product of an *Arabidopsis* gene. This gene product is currently being characterised (F. Nogu  , personal communication). FasV is part of the *fas* operon of *Rhodococcus fascians* encoding genes required for fasciation in plants (Crespi *et al.*, 1994). One of the genes belonging to this operon codes for cytokinin synthase (Crespi *et al.*, 1992). The FasV protein has been tentatively described as an electron transporter by Goethals *et al.* (1995). FasV homology with FAD-dependent oxidases was already pointed out by Mushegian and Koonin (1995). This homology indicates that maize CKO is most likely a flavoprotein. This conclusion is further supported by our observation (results not shown) that maize CKO is strongly inhibited by diphenyliodonium (DPI), an inhibitor of flavoprotein oxidoreductases (O'Donnell *et al.*, 1993).

By analysing the activity of glycosylated and unglycosylated CKOs in *Phaseolus* callus tissues, Kaminek and Armstrong (1990) suggested that a compartmentation could exist in the cells that keep the glycosylated form in the cell wall or plasmalemma and the unglycosylated form in an internal compartment. Is glycosylation important for functionality and/or localization? This question requires further investigation. Because of our purification scheme, we isolated a glycosylated protein, and Southern analysis indicates that the gene is likely to be present in one copy. We cannot rule out, however, the possibility that some related genes exist. Moreover, at least for the four

sequenced peptides, no difference was observed between the commercial sweet corn variety and the hard corn field variety Nobilis. We did not isolate any cDNA other than the sequence reported here. However, CKO may belong to a family of genes with different expression patterns as suggested by Jones and Schreiber (1997). According to Northern analysis, the CKO gene seems to be weakly expressed in kernels harvested about two weeks after anthesis, a time that has been reported to correspond with a peak of CKO activity (Dietrich *et al.*, 1995).

We now possess a tool for analysing the role of cytokinin oxidase in plant development in general and in endosperm/embryo development in particular. It should be interesting to overexpress CKO in plants and to identify if, and how, this overexpression affects their phenotypes. Transgenic plants overproducing cytokinins have been already produced by transformation with the *Agrobacterium tumefaciens IPT* gene (Faiss *et al.*, 1997; Smigocki and Owens, 1989). Interestingly, increased cytokinin metabolism/inactivation seemed to occur in these plants. Zhang *et al.* (1995) observed an NAA-mediated oxidative breakdown of both exogenous radiolabeled zeatin-type cytokinins and endogenous cytokinins in tobacco tissue transformed with the *IPT* gene. These observations were supported by Motyka *et al.* (1996) who showed an increase in cytokinin oxidase activity in both leaves and roots of transgenic tobacco plants expressing the *IPT* gene. It will be of special interest to confirm if and how cytokinin homeostasis is affected in plants overexpressing CKO or in mutants of *Arabidopsis thaliana* having a disrupted CKO gene.

## Experimental procedures

### Plant material

Commercial sweet corn was used for protein purification. For molecular biology experiments, kernels originated from field hard corn variety Nobilis, a mid-early hybrid, Pau Semences, 64230 Lescar, France. Cobs were cleaned, frozen in liquid nitrogen and kept at  $-70^{\circ}\text{C}$  before use. The Gransden wild-type strain of the moss *Physcomitrella patens* (Ashton and Cove, 1977) was used for transient expression assays.

### Chemicals

[ $^3\text{H}$ ]-iPA ( $17.8\text{ mCi } \mu\text{mol}^{-1}$ ) was prepared according to Laloue and Fox (1989). HA8 was a gift from R. Mornet (Angers University, France). iP, iPA, methyl- $\alpha$ -D-mannopyranoside, 2,5-diphenyloxazole, trypsin inhibitor, leupeptin, DTT, bovine serum albumin, Chaps, and the DOC-TCA kit were purchased from Sigma, St Louis, MO, USA. Pefabloc was from Interchim, Montluçon, France; concanavalinA-Sepharose was from Pharmacia, Uppsala, Sweden; polyclar AT was from Boehringer, Ingelheim, Heidelberg, Germany.

### CKO purification

**Extraction.** Eight aliquots of 50 g frozen kernels were ground to a fine powder in a coffee grinder with four pulses of 10 sec each.

The resulting powder was homogenized in a cold mortar with 800 ml of cold extraction buffer (0.1 M phosphate buffer pH 6.8 adjusted prior to use at  $25\text{ ng l}^{-1}$  trypsin inhibitor,  $1\text{ mg l}^{-1}$  leupeptin,  $1\text{ mM}$  Pefabloc,  $2\text{ mM}$  DTT,  $5\text{ mM}$  EDTA,  $33\text{ g l}^{-1}$  Polyclar AT). Homogenates were pooled, left at  $4^{\circ}\text{C}$  for 30 min with intermittent stirring and centrifuged at  $20\,000\text{ g}$  for 90 min. Supernatants were pooled and filtrated through two layers of Miracloth (Calbiochem, La Jolla, CA, USA).

**Ammonium sulfate fractionation.** Solid ammonium sulfate was added to the extract to 30% saturation over a period of 20 min at  $4^{\circ}\text{C}$ . After 1 h, the precipitate was eliminated by centrifugation ( $20\,000\text{ g}$  for 90 min). Ammonium sulfate was again added in the same way to the collected supernatant to 50% saturation. After centrifugation, the pellet was then solubilized in  $2\text{ mM}$  phosphate buffer pH 6.8 containing  $2\text{ mM}$  DTT and the solution was dialysed overnight against three changes of EB buffer ( $20\text{ mM}$  Tris-HCl buffer pH 7.4,  $0.5\text{ M}$  NaCl and  $2\text{ mM}$  DTT). The dialysed extract was clarified by centrifugation at  $40\,000\text{ g}$  for 1 h, aliquoted and stored at  $-20^{\circ}\text{C}$ .

**ConcanavalinA-Sepharose chromatography.** The extract was applied on a concanavalinA-Sepharose column (2.5 cm inner diameter, 6 cm height) equilibrated with  $100\text{ ml}$  EB buffer. The sample ( $65\text{ ml}$  for  $400\text{ g}$  kernels) was applied to the column which was washed in EB buffer without NaCl until the return of OD to the baseline and eluted in the same buffer containing  $0.2\text{ M}$  methyl- $\alpha$ -D-mannopyranoside. Four ml fractions were collected and assayed for CKO activity. The appropriate fractions were pooled, aliquoted and stored at  $-20^{\circ}\text{C}$ .

**Anion exchange chromatography.** An FPLC apparatus (Pharmacia) equipped with an anion exchange column (ResourceQ, Pharmacia, 6 ml) was used. Equilibration was performed in T buffer ( $20\text{ mM}$  Tris-HCl buffer pH 8.0). Two ml sample aliquots were applied to the column. The column was then washed in T buffer and eluted in a linear gradient of  $1\text{ M}$  NaCl in the same buffer (increments were  $2.5\text{ mM NaCl min}^{-1}$  for 1 h). Two ml fractions were collected and assayed for CKO activity and/or radioactivity as required. The appropriate fractions were pooled, aliquoted and stored at  $-20^{\circ}\text{C}$ .

### Gel permeation chromatography

A TSK3000SW gel permeation column (Tosohas, Montgomery Ville, PA, USA) ( $7.5\text{ mm I.D.} \times 300\text{ mm height}$ ) was equilibrated with Tris-HCl  $20\text{ mM}$  pH 7.8,  $\text{Na}_2\text{SO}_4$   $0.1\text{ M}$ ,  $\text{Na}_2\text{EDTA}$   $5\text{ mM}$  buffer. Elution was performed in the same buffer. The gel filtration standards were from BioRad Laboratories, Hercules, CA, USA. Fractions of  $250\text{ } \mu\text{l}$  were collected and assayed for CKO activity and radioactivity.

### Assay for CKO activity

Unless otherwise stated, assays were performed according to Laloue and Fox (1989), using [ $^3\text{H}$ ]-iPA  $2\text{ } \mu\text{M}$ . The radioactive adenosine formed was separated from the substrate by HPLC on a C<sub>8</sub> Lichrospher 60, RPselect B,  $5\text{ } \mu\text{m}$  column ( $125 \times 4\text{ mm}$ ) from Merck. The HPLC system was a Waters gradient instrument 600E equipped with a programmable multiwavelength detector 490E. Radioactivity was measured with an on line Flo-one  $\beta$  instrument from Radiomatic.

### Photoaffinity labelling

One ml protein aliquots obtained after conA-chromatography were incubated with 0.5  $\mu\text{M}$  [ $^3\text{H}$ ]-azidoCPPU (333 Gbq  $\text{mmol}^{-1}$ ) at 4°C for 30 min (unless otherwise stated) on a shaker in the dark. After transfer in quartz cuvettes, samples were placed at a distance of 25 cm from a 254 nm, 6 W lamp (Bioblock Scientific, Illkirch, France) and irradiated for 5 min (energy = 0.3  $\text{mW cm}^{-2}$ ). Photoincorporation of cold azidoCPPU was performed five times in the presence of 10  $\mu\text{M}$  azidoCPPU, prior to preparative 2D-gel electrophoresis.

Before anionic exchange and gel permeation chromatography, excess of radioactive probe was removed by using desalting columns (NAP columns from Pharmacia) equilibrated with T buffer.

### Polyacrylamide gel electrophoresis

**Sample preparation.** After anionic exchange chromatography, proteins were precipitated with DOC-TCA, pelleted by centrifugation at 12 000  $g$  for 15 min, washed three times in methanol containing 0.1 M ammonium acetate and once in methanol, vacuum-dried and then solubilized in Laemmli (1970) or IEF buffer (8 M urea, 2% Phosphoryltes 3-10, 32 mM DTT, 8 mM Chaps, 10  $\mu\text{M}$  leupeptin, 1 mM Pefabloc and a few bromophenol blue crystals) for 1 and 2D-gel electrophoresis, respectively.

**SDS-PAGE and fluorography.** SDS-PAGE was performed on 0.75 mm thick 12% polyacrylamide running gel with 5% polyacrylamide stacking gel. The peptide pattern was visualized by Coomassie blue or silver staining. [ $^{14}\text{C}$ ]-labelled molecular weight markers were from Amersham, Cleveland, OH, USA.

For fluorographic analysis, the Coomassie blue stained gels were soaked in glacial acetic acid containing 20% (w/v) 2,5-diphenyloxazole for 60 min according to Skinner and Griswold (1983). Radiolabelling was visualized after exposure at -70°C of the dried gels to preflashed Kodak X-O-Mat films. The fluorographic signals were quantitated after digitalization with an OmniMedia 6CX-XRS scanner and analysis of the data on a Bioimage 1D software (Millipore Co, Bedford, MA, USA). They were corrected to take into account differences in the protein contents of the lanes which were measured on the Coomassie blue stained gel by the same technique.

**2D-gel electrophoresis.** Samples were submitted to 2D-gel electrophoresis on a Multiphor II apparatus from Pharmacia. IEF was performed on 11 cm or 18 cm Immobiline Dry strips pH 4-7 for 16 h at 33 or 44-45 kVh, respectively, and, if necessary, stored at -70°C before use. For the second dimension, the IEF gel strips were soaked at room temperature in 50 mM Tris-HCl buffer pH 6.8, 6 M urea, 30% glycerol, 1% SDS (w/v) for 15 min in the presence of 16 mM DTT and for 15 additional min in the presence of 240 mM iodoacetamide. The strips were then transferred onto the surface of a horizontal ExcelGel<sup>TM</sup> (Pharmacia, 8-18% gradient SDS-polyacrylamide) according to the procedure described by Pharmacia. Electrophoresis was carried out at 20 mA in the stacking gel and 40 mA in the running gel. Peptides were visualised by silver staining for analytical electrophoresis, amido-black for preparative gels and Coomassie blue for fluorography. In that case, gels were removed from their gel-bond support and treated as SDS-PAGE gels.

### Protein determination

All protein determinations were done according to the Lowry's procedure with the Protein Assay kit from Sigma. Bovine serum albumin was used as a standard.

### Determination of amino acid sequences

Spots were cut out of the amido-black stained 2D-gels and partially dehydrated in a Speed-Vac. Gel pieces were rehydrated in 150-200  $\mu\text{l}$  of 0.1 M Tris-HCl, pH 8.6, 0.01% Tween 20 (Pierce, Rockford, IL, USA) and digested with endoproteinase Lys-C from *Lysobacter enzymogenes* (Boehringer, Mannheim, Germany) at a final concentration of 2  $\mu\text{g ml}^{-1}$  for 18 h at 35°C.

The supernatant was recovered and the pellet was rinsed with 60% acetonitrile. The acetonitrile rinse was added to the supernatant and acetonitrile was removed in a Speed-Vac. Sample was injected onto a DEAE-HPLC column linked to a C18 reverse phase HPLC column eluted with a 0-45% acetonitrile, 0.1% TFA gradient (Kawasaki and Suzuki, 1990). Peaks recorded at 210 nm were collected manually and frozen (-20°C) until sequencing. Sequencing was performed on Applied Biosystems 473 and 494 sequencers.

### Nucleic acid isolation

Plasmid DNA was purified on QIAGEN-tips (QIAGEN GmbH, Hilden, Germany). Maize DNA was prepared from kernels according to Dellaporta *et al.* (1983) and purified by CsCl gradient. Total RNA extraction from kernels was adapted from Hall *et al.* (1978). Poly A<sup>+</sup> RNA was purified by two rounds of spun-column chromatography by using the mRNA purification kit of Pharmacia.

### DNA cloning

DNA amplification fragments were excised from electrophoresis gels and recovered by using a GeneClean kit (Bio 101 Inc, Vista, CA, USA). They were cloned in a pBluescript (KS<sup>+</sup>) T-vector. This vector was prepared by digesting the plasmid with *EcoRV* restriction enzyme and adding thymidine 3'-overhangs according to Marchuk *et al.* (1991).

The plasmids were used to transform XL1-Blue (Stratagene, La Jolla, Ca, USA) or DH10B (Gibco BRL, Paisley, Scotland) *Escherichia coli* strains. Oligonucleotides spanning the start and the stop codon DNA region (CKs8, CKa8) were used to amplify the ORF-cDNA. Amplification was performed with the Advantage Klen Taq Polymerase (Clontech Laboratories, Palo Alto, CA, USA) in the presence of DMSO 4%. One pBluescript clone in which the CKO-ORF was in opposite orientation compared to the  $\beta$ -gal promoter (clone 16) was completely sequenced. The CKO-ORF fragment was directionally cloned in the plant vector pLBR19, a derivative of pJIT60 (Guerineau *et al.*, 1992). This vector carries a duplicated CaMV 35S promoter and a CaMV terminator. A *SalI* blunt-ended *EcoRI* fragment was cloned in pLBR19 that had been digested with *HindIII*, made blunt and further digested with *SalI*. The junction between the promoter and the ORF was checked by sequencing. This construct was named pLBR1916. pLBR19 with GUS-ORF was a gift of P. Mourrain (INRA Versailles, France).

### DNA sequencing

DNA was sequenced by Sanger's method. Cloned fragments were either sequenced in a cycle sequencing procedure with the Dye primer kit (PRISM<sup>TM</sup>, Perkin Elmer, New Jersey, USA) or with the Dye terminator kit (ABI PRISM<sup>TM</sup>, Perkin Elmer) and specific internal primers (CKs10, CKs11, CKs12, CKs13, CKa16, CKa10, CKa11, CKa12, CKa13). Electroeluted PCR fragments were sequenced with the Dye terminator kit (ABI PRISM<sup>TM</sup>, Perkin Elmer). Analysis was performed on an Applied Biosystems 370A

DNA sequencer in the laboratory or at Cybergène (Saint-Malo, France).

### Southern analysis

Genomic DNA (8 µg) was fractionated on a 0.8% agarose gel and blotted onto Hybond™ N<sup>+</sup> membranes. The blot was fixed by UV crosslinking in an UV Stratalinker (Stratagene). Membrane was probed with gel-purified restriction fragments that were radiolabelled with the hexamer kit of Pharmacia. Probes were a 1.1 kb *NotI*-*Bam*HI internal restriction fragment of a 5'-RACE clone and a 0.6 kb *SacI*-*Hind*III restriction fragment of clone 16, respectively. Pre-hybridisation and hybridisation steps were performed at 65°C in formamide-free Church's buffer (0.25 M Na<sub>2</sub>HPO<sub>4</sub> pH 7.4, 2 mM EDTA, 7% SDS, 0.2 mg ml<sup>-1</sup> heparin and 0.1 mg ml<sup>-1</sup> denatured salmon sperm DNA). Washing was done for 15 min in 2 × SSC, 0.1% SDS at 50°C (i.e. low stringency conditions) and the blot was submitted to autoradiography. Membrane dehybridisation was performed according to the manufacturer's instructions and total removal of the probe was checked by autoradiography. Pre-hybridisation and hybridisation steps were performed as described above. Washing was done for 30 min in 2 × SSC, 0.1% SDS at 55°C followed by 30 min washing in 0.2 × SSC, 0.1% SDS at 65°C (i.e. high stringency conditions).

### Northern analysis

Total RNA (10 µg) and poly A<sup>+</sup> RNA (1 µg) were fractionated on a 1.2% agarose formaldehyde gel and blotted onto a GeneScreen membrane (DuPont-NEN Research Products, Boston, MA, USA). RNA immobilisation and hybridisation were identical to Southern analysis. Washing was done for 20 min in 2 × SSC, 0.1% SDS at 55°C. Molecular weight markers (0.24–9.5 kb) were from Gibco BRL.

### Oligonucleotide sequences

#### (a) Degenerate oligonucleotides

Coa1: CAT(TC)TAC(AG)AAIC(TG)(GA)TTCCA = an antisense primer related to peptide 1 (WNRFFVEM);  
Coa23: TT(AG)AAIAT(AG)TC(TC)TG(ACGT)CC(ACGT)GG = an antisense primer related to the C-terminus (PGQDIFN) of peptide 2;  
Co51: AT(CT)TICA(AG)GGIACIGA(CT)AT(CT)GT(ACGT)GG = a sense primer related to the N-terminus (ILQGTDIVG) of peptide 5;  
Co52: GTIGGIC(CT)TIAT(CT)GTITA(CT)CC = a nested sense primer related to peptide 5 (VGPLIVYP).

#### (b) Specific oligonucleotides

CKs2: TGGCGCCCAACGACCTG; CKs3: TGCAGGAGCAGAACAGG;  
CKs6: ACAATCCATGTGGGACGACGGCA;  
CKs7: GGATCTGCGCTTCTGCGACCTC;  
CKs8: GCTAATGGCGGTGGTTTATTACCTGCTG;  
CKs10: CCTACACCATCGCGTTCC; CKs11: TTCCTTGACCGGGTGCAGG;  
CKs12: TCCTGTACCCGTCGTCCAC; CKs13: ATGGGGAGATGGT-GACGTG;  
CKs16: CGACCGACCTGGCGAACAC; CKa5: GCCGCGCCAGGTA-GGTCTTG;  
CKa7: GGACCCAGTCACTGCGGTCCGTGT;  
CKa8: CCAAGAACTAAAAACATGCATGGGCTAT;  
CKa10: CTGTGTGAGCGCCACCTCC; CKa11: CTGGCCCATGAGGG-AGTGG;

CKa12: CACGTAGCTCATCGGGCC; CKa13: GATCCACACCTG-CTCGCC.

Oligonucleotides were purchased from Genset, Paris, France and Eurogentec, Liège, Belgium.

### RT-PCR

The cDNA first strand was reverse transcribed from 2 µg poly A<sup>+</sup> RNA by using the Pharmacia Ready to Go kit and random hexamers as primers (Boehringer). Six independent reactions were performed. Each reaction product was split and used for two RT-PCR reactions. A touch-down procedure (Don *et al.*, 1991) was used, consisting of 11 cycles with an annealing temperature scaled down from 1°C each cycle, starting from 62°C to 52°C (45 sec at 94°C, 30 sec at the annealing temperature and 2 min at 72°C) followed by 35 cycles (45 sec at 94°C, 30 sec at 52°C and 1 min 30 sec at 72°C). The PCR reaction mixture was as advised by the company. Forty picomoles of the degenerate oligonucleotides (Coa1, Coa23, Co51) were used in a 50 µl reaction. Bacterial transformants were screened by PCR with Co52 and Coa1.

### RACE PCR

Poly A<sup>+</sup> RNA was treated with the Marathon™ cDNA amplification kit from Clontech. One µg was used in each reaction. The procedure was carried out exactly as advised by the company. Specific primers were CKs6, CKs7 and CKa7.

The final 3'-RACE reaction product was obtained after two rounds of amplification with the Expand™ Long template PCR system (Boehringer) and primers CKs6-AP1 followed by a nested PCR with primers CKs7-AP2 and the Taq polymerase (Perkin Elmer), respectively. AP1 and AP2 are the anchored adaptors included in the Marathon™ kit. A 5'-RACE reaction product was visible after one round of amplification with the Clontech Advantage Klen Taq Polymerase in the presence of DMSO 4%. It was reamplified by gel puncturing by using the same primers. Bacterial transformants were screened by PCR with CKs2, CKs3, CKa5 and the universal primers.

### Transient expression in moss protoplasts

Moss cultures and protoplast preparations were performed according to Schaefer and Zryd (1997). Improved protoplast preparations were obtained by overnight digestion at 25°C in the dark with 0.05% driselase (Fluka Chemie AG, Switzerland) and 0.02% macerozyme R10 and 0.1% cellulase Onozuka R10 (Yakult Biochemicals, Nishinomiya, Japan) according to Bourgin *et al.* (1979). Briefly, after several washes in mannitol 8.5%, the protoplasts were counted and resuspended in mannitol 8.5%, MgCl<sub>2</sub> 15 mM, MES 0.1%, pH 5.6. PEG-mediated direct DNA transfer into protoplasts was performed as described by Schaefer *et al.* (1994). The samples were kept for 20 h in the dark and transferred for 24 h in the light. The culture medium was collected after centrifugation. Aliquots were analysed for CKO activity as described, except that [<sup>3</sup>H]-IPA was 0.05 µM. Cells were washed twice in new culture medium and disrupted by three cycles of freezing and thawing. The resulting crude extracts were assayed for CKO and GUS activities. GUS assay was performed according to Jefferson *et al.* (1987).

### Acknowledgements

We thank Fabien Nogué for his help in the completion of the manuscript. We also thank Herman Höfte and Ian Small for

critically reading the manuscript. The technical help of Amel Majira is acknowledged. We thank Daniel Duval for providing the maize line, Jean-Pascal Meunier and Jean-Marie Pollien for growing the plants and Kirk Schnorr for his past contribution to this project. This work is dedicated to the memory of Jean-Pierre Bourgin on the occasion of the 4th anniversary of his death.

## References

- Armstrong, D.J. (1994) Cytokinin oxidase and the regulation of cytokinin degradation. In *Cytokinins, Chemistry, Activity and Function* (Mok, D.W.S. and Mok, M.C., eds). Boca Raton: CRC Press, pp. 139–154.
- Ashton, N.W. and Cove, D.J. (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants of the moss, *Physcomitrella patens*. *Mol. Gen. Genet.* **154**, 87–95.
- Ashton, N.W., Grimsley, N.H. and Cove, D.J. (1979) Analysis of gametophytic development in the moss, *Physcomitrella patens*, using auxin and cytokinin resistant mutants. *Planta*, **144**, 427–435.
- Bourgin, J.P., Chupeau, Y. and Missonier, C. (1979) Plant regeneration from mesophyll protoplasts of several *Nicotiana* species. *Physiol. Plant.* **45**, 288–292.
- Brzobohaty, B., Moore, I., Kristoffersen, P., Bako, L., Campos, N., Schell, J. and Palme, K. (1993) Release of active cytokinin by a  $\beta$ -glucosidase localized to the maize root meristem. *Science*, **262**, 1051–1054.
- Burch, L.R. and Horgan, R. (1989) The purification of cytokinin oxidase from *Zea mays* kernels. *Phytochemistry*, **28**, 1313–1319.
- Burch, L.R. and Horgan, R. (1992) Cytokinin oxidase and the degradative metabolism of cytokinins. In *Physiology and Biochemistry of Cytokinins in Plants* (Kaminek, M., Mok, D.W.S. and Zamilova, E., eds). The Hague: SPB Academic Publishing, pp. 29–32.
- Chatfield, J.M. and Armstrong, D.J. (1986) Regulation of cytokinin oxidase activity in callus tissues of *Phaseolus vulgaris* L. *Cv Great Northern*. *Plant Physiol.* **80**, 493–499.
- Chatfield, J.M. and Armstrong, D.J. (1988) Cytokinin oxidase from *Phaseolus vulgaris* callus cultures: affinity for concanavalin A. *Plant Physiol.* **88**, 245–247.
- Chen, C.M. (1981) Biosynthesis and enzymic regulation of the interconversion of cytokinin. In *Metabolism and Molecular Activities of Cytokinins* (Guern, J. and Péaud-Lenoël, C., eds). Berlin: Springer-Verlag, pp. 34–43.
- Crespi, M., Messens, E., Caplan, A.B., van Montagu, M. and Desomer, J. (1992) Fasciation induction by the phytopathogen *Rhodococcus fascians* depends upon a linear plasmid encoding a cytokinin synthase gene. *EMBO J.* **11**, 795–804.
- Crespi, M., Verecke, D., Temmerman, W., van Montagu, M. and Desomer, J. (1994) The *fas* operon of *Rhodococcus fascians* encodes new genes required for efficient fasciation of host plants. *J. Bact.* **176**, 2492–2501.
- Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) A plant DNA miniprep preparation Version II. *Plant Mol. Biol. Report*, **1**, 19–21.
- Dias, M., Mornet, R. and Laloue, M. (1995) Synthesis, azido-tetrazole equilibrium studies and biological activity of 1-(2-azido-6-chloro-pyrid-4-yl)-3-phenylurea, a photoaffinity labeling reagent for cytokinin-binding proteins. *Bioorg. Medic. Chem.* **3**, 361–366.
- Dietrich, J.T., Kaminek, M., Blevins, D.G., Reinbott, T.M. and Morris, R.O. (1995) Changes in cytokinins and cytokinin oxidase activity in developing maize kernels and the effects of exogenous cytokinin on kernel development. *Plant Physiol. Biochem.* **33**, 327–336.
- Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. and Mattick, J.S. (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucl. Acids Res.* **19**, 4008.
- Faiss, M., Zalubilova, J., Strnad, M. and Schmülling, T. (1997) Conditional transgenic expression of the *ipt* gene indicates a function for cytokinins in paracrine signaling in whole tobacco plants. *Plant J.* **12**, 401–415.
- Goethals, K., Verecke, D., Temmerman, W., Maes, T., Kalkus, J., Simon-Mateo, C. and Van Montagu, M. (1995) Cytokinin production by the phytopathogenic bacterium *Rhodococcus fascians*. *Med. Fac. Landbouww. University of Gent* **60/4a**, 1553–1558.
- Gonneau, M., Mornet, R. and Laloue, M. (1998) A *Nicotiana plumbaginifolia* protein labeled with an azido cytokinin agonist is a glutathione S-transferase. *Physiol. Plant.* **103**, 114–124.
- Guerineau, F., Lucy, A. and Mullineaux, P. (1992) Effect of two consensus sequences preceding the translation initiator codon on gene expression in plant protoplasts. *Plant Mol. Biol.* **18**, 815–818.
- Hall, T.C., Ma, Y., Buchbinder, B.U., Pyne, J.W., Sun, S.M. and Bliss, F.A. (1978) Messenger RNA for G1 protein of French bean seeds: Cell-free translation and product characterization. *Proc. Natl Acad. Sci. USA*, **75**, 3196–3200.
- Hecht, S.M., Frye, R.B., Werner, D., Hawrelak, S.D., Skoog, F. and Schmitz, R.Y. (1975) On the 'activation' of cytokinins. *J. Biol. Chem.* **250**, 7343–7351.
- Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions:  $\beta$ -glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.
- Jones, R.J. and Schreiber, B.M.N. (1997) Role and function of cytokinin oxidase in plants. *Plant Growth Regul.* **23**, 123–134.
- Kaminek, M. and Armstrong, D.J. (1990) Genotypic variation in cytokinin oxidase from *Phaseolus* callus cultures. *Plant Physiol.* **93**, 1530–1538.
- Kawasaki, H. and Suzuki, K. (1990) Separation of peptides dissolved in a sodium dodecyl sulfate solution by reversed-phase liquid chromatography: Removal of sodium dodecyl sulfate from peptides using an ion-exchange precolumn. *Anal. Biochem.* **186**, 264–268.
- Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
- Laloue, M. (1980) Etude de l'importance fonctionnelle des processus d'absorption et des transformations métaboliques des cytokinines dans l'expression de leur activité biologique chez les cellules de *Nicotiana tabacum*. Thèse, de Doctorat d'Etat ès-sciences naturelles. Université Pierre et Marie Curie, Paris.
- Laloue, M. and Fox, J.E. (1985) Characterization of an imine intermediate in the degradation of isopentenylated cytokinins by a cytokinin oxidase from wheat. In *Abstracts 12th International Conference on Plant Growth Substances*. Heidelberg: International Plant Growth Association.
- Laloue, M. and Fox, J.E. (1989) Cytokinin oxidase from wheat. Partial purification and general properties. *Plant Physiol.* **90**, 899–906.
- Marchuk, D., Drumm, M., Saulino, A. and Collins, F.S. (1991) Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* **19**, 1154.
- Martin, R.C., Mok, M.C. and Mok, D.W.S. (1997) Protein processing and auxin response in transgenic tobacco harboring a putative cDNA of zeatin 0-xylosyltransferase from *Phaseolus vulgaris*. *Plant J.* **12**, 305–312.
- McGaw, B.A. and Burch, L.R. (1995) Cytokinin biosynthesis and metabolism. In *Plant Hormones – Physiology, Biochemistry and Molecular Biology* (Davies, P.J., ed.). Dordrecht: Kluwer Academic Publishers, pp. 318–339.

- Meilan, R. and Morris, R.O. (1994) Cloning the cytokinin oxidase gene. *Plant Physiol.* **105** (Suppl.), 68.
- Miller, C.O., Skoog, F., Okomura, F.S., von Saltza, M.H. and Strong, F.M. (1956) Isolation, structure and synthesis of kinetin, a substance promoting cell division. *J. Am. Chem. Soc.* **78**, 1345–1350.
- Moffatt, B.A., McWhinnie, E.A., Burkhardt, W.E., Pasternak, J.J. and Rothstein, S.J. (1992) A complete cDNA for adenine phosphoribosyltransferase from *Arabidopsis thaliana*. *Plant Mol. Biol.* **18**, 653–662.
- Mok, D.W.S. and Martin, R.C. (1994) Cytokinin metabolic enzymes. In *Cytokinins, Chemistry, Activity and Function* (Mok, D.W.S. and Mok, M.C., eds). Boca Raton: CRC Press, pp. 129–137.
- Morris, R.O. (1995) Genes specifying auxin and cytokinin biosynthesis in prokaryotes. In *Plant Hormones – Physiology, Biochemistry and Molecular Biology* (Davies, P.J., ed.). Dordrecht: Kluwer Academic Publishers, pp. 318–339.
- Motyka, V., Faiss, M., Strnad, M., Kaminek, M. and Schmülling, T. (1996) Change in cytokinin content and cytokinin oxidase activity in response to derepression of *ipt* gene transcription in transgenic tobacco calli and plants. *Plant Physiol.* **112**, 1035–1043.
- Mushegian, A.R. and Koonin, E.V. (1995) A putative FAD-binding domain in a distinct group of oxidases including a protein involved in plant development. *Prot. Sci.* **4**, 1243–1244.
- Nogué, F., Mornet, R. and Laloue, M. (1996) Specific photoaffinity labeling of a thylakoid membrane protein with an azido-cytokinin agonist. *Plant Growth Reg.* **18**, 51–58.
- O'Donnell, V.B., Tew, D.G., Jones, O.T.G. and England, P.J. (1993) Studies on the inhibitory mechanism of iodonium compounds with special reference to neutrophil NADPH oxidase. *Biochem. J.* **290**, 41–49.
- Schaefer, D.G., Bisztray, G. and Zrýd, J.-P. (1994) Genetic transformation of the moss *Physcomitrella patens*. In *Biotechnology in Agriculture and Forestry* Volume 29 (Bajaj, Y.P.S., ed.). Berlin: Springer-Verlag, pp. 349–364.
- Schaefer, D. and Zrýd, J.-P. (1997) Efficient gene targeting in the moss *Physcomitrella patens*. *Plant J.* **11**, 1195–1206.
- Schnorr, K.M., Gaillard, C., Biget, E., Nygaard, P. and Laloue, M. (1996) A second form of adenine phosphoribosyltransferase in *Arabidopsis thaliana* with relative specificity towards cytokinins. *Plant J.* **9**, 891–898.
- Schreiber, B.M.N., Roessler, J.A. and Jones, R.J. (1995) Polyclonal antibodies to maize cytokinin oxidase. *Plant Physiol.* **108** (Suppl.), 80.
- von Schwartzberg, K., Kruse, S., Reski, R., Moffatt, B. and Laloue, M. (1998) Cloning and characterization of an adenosine kinase from *Physcomitrella* involved in cytokinin metabolism. *Plant J.* **13**, 249–257.
- Skinner, M.K. and Griswold, M.D. (1983) Fluorographic detection of radioactivity in polyacrylamide gels with 2,5-diphenyloxazole in acetic acid and its comparison with existing procedures. *Biochem. J.* **209**, 281–284.
- Skoog, F. and Miller, C.O. (1957) Chemical regulation of growth and organ formation in plant tissues cultured in vitro. *Soc. Exp. Biol.* **11**, 118–131.
- Smigocki, A.C. and Owens, L.D. (1989) Cytokinin-to-auxin ratios and morphology of shoots and tissues transformed by a chimeric isopentenyl transferase gene. *Plant Physiol.* **91**, 808–811.
- Terrine, C. and Laloue, M. (1980) Kinetics of N<sup>6</sup>-( $\Delta^2$ -isopentenyl) adenosine degradation in tobacco cells. Evidence of a regulatory mechanism under the control of cytokinins. *Plant Physiol.* **65**, 1090–1095.
- Walsh, C. (1982) Suicide substrates: mechanism-based enzyme inactivators. *Tetrahedron*, **38**, 871–909.
- Wang, J. and Letham, D.S. (1995) Cytokinin oxidase-purification by affinity chromatography and activation by caffeic acid. *Plant Sci.* **112**, 161–166.
- Whitty, C.D. and Hall, R.H. (1974) A cytokinin oxidase in *Zea mays*. *Can. J. Biochem.* **52**, 789–799.
- Zhang, R., Zhang, X., Wang, J., Letham, D.S., McKinney, S.A. and Higgins, T.J.V. (1995) The effect of auxin on cytokinin levels and metabolism in transgenic tobacco tissue expressing an *ipt* gene. *Planta*, **196**, 84–94.

EMBL Data Library accession number Y18377.

# Note Added in Proof

During the completion of this paper, the sequence of the maize *ckx1* gene (AF044603) from R.O. Morris and J.G. Laskey (University of Missouri, Columbia, USA) was made available in the databank. Our cDNA sequence differs at seven positions in the ORF leading to three amino acid changes.

# Isolation of a Gene Encoding a Glycosylated Cytokinin Oxidase from Maize<sup>1</sup>

Roy O. Morris,<sup>\*,2</sup> Kristin D. Bilyeu,<sup>\*</sup> James G. Laskey,<sup>\*</sup> and Nordine N. Cheikh<sup>†</sup>

<sup>\*</sup>Biochemistry Department, 117 Schweitzer Hall, University of Missouri, Columbia, Missouri 65211; and

<sup>†</sup>Monsanto Company, 700 Chesterfield Parkway North, St. Louis, Missouri 63198

Received December 28, 1998

The major cytokinin oxidase in immature maize kernels was purified to homogeneity. Selected tryptic peptides were used to design degenerate oligonucleotide primers for PCR isolation of a fragment of the oxidase gene. Hybridization of the PCR fragment to a maize genomic library allowed isolation of a full-length cytokinin oxidase gene (*ckx1*). The gene encodes a protein of approximately 57 kDa that possesses a signal peptide, eight consensus *N*-glycosylation sequences and a consensus FAD binding sequence. Expression of *ckx1* in *Pichia* caused secretion of active glycosylated cytokinin oxidase that contains a substrate-reducible FAD. The gene displays sequence homology with a putative oxidoreductase from *Arabidopsis thaliana* and with the *fas5* gene from *Rhodococcus fascians*. © 1999 Academic Press

**Key Words:** cytokinin oxidase; glycoprotein; maize; *Zea mays*; *Pichia pastoris*; PCR.

The enzymes responsible for cytokinin biosynthesis and catabolism in plants have proved remarkably recalcitrant to purification and characterization. For example, no structural gene encoding cytokinin biosynthesis has yet been reported from a higher plant, although homologous genes (*ipt*, *tzs*, *etz*) encoding cytokinin biosynthetic prenyl transferases have been characterized from a number of gall-forming plant pathogenic bacteria (1–3). Likewise, plant genes encoding cytokinin catabolic enzymes have eluded characterization. Of the cytokinin catabolic enzymes, cytokinin oxidases are ubiquitous and probably of greatest sig-

nificance (4,5). However, purification of these oxidases has proved challenging because glycosylated and non-glycosylated oxidases are present in most plants (6) and because both appear to be present at low levels. In spite of these problems, the maize and wheat glycosylated cytokinin oxidases have been purified to near homogeneity (7,8) but the genes have not been isolated.

We now report the purification of the predominant cytokinin oxidase from immature maize kernels and the isolation and expression of its structural gene (*ckx1*). It is a glycosylated flavoprotein oxidoreductase with homology to an uncharacterized oxidoreductase located on chromosome II of *Arabidopsis* (9) and to the *fas5* gene product of *Rhodococcus fascians* (10).

## MATERIALS AND METHODS

**Enzyme assays.** Cytokinin oxidase activity was measured either by the method of Libreros-Minotta and Tipton (1995) or by a continuous dichlorophenolindophenol (DCPIP) reduction assay (11).

**Enzyme extraction.** Immature kernels were harvested from field-grown maize (Pioneer 3180) five to eight days after pollination and frozen at –80°C. Frozen powdered kernels (1 Kg aliquots) were blended with 1700 ml of a buffer containing 50 mM Tris-Cl, 5 mM EDTA, 0.4% (w/v) ascorbic acid, and 10 mM  $\beta$ -mercaptoethanol, pH 8.5. An excess of acid-washed solid polyvinylpyrrolidone (PVPP) was stirred in. After filtration and centrifugation to remove debris, polyethyleneimine was added to 0.05% (v/v) and, after recentrifugation (23,500  $\times$  g for 10 minutes), the supernatant was filtered through a pad of PVPP and the 40–65% ammonium sulfate fraction was collected. It was dissolved in a buffer containing 10 mM Tris-Cl and 1 mM EDTA, pH 8.5 containing 1 mM  $\beta$ -mercaptoethanol and stored at –80°C. Aliquots were pooled and fractionated by preparative scale (500 ml) DEAE-cellulose chromatography in this buffer (without mercaptoethanol) and a gradient of KCl. The major oxidase peak was applied to a concanavalin A-agarose column in a buffer containing 20 mM Tris-Cl, 0.5 M NaCl, and 1.0 mM CaCl<sub>2</sub>, pH 7.4. Glycosylated proteins were eluted with a step gradient of 1 M  $\alpha$ -D-methylmannoside. The resulting glycosylated oxidase fraction was purified further by anion exchange chromatography (FPLC MonoQ, 1 ml, Pharmacia, Piscataway, NJ, USA) and by hydrophobic interaction chromatography (Phenyl superose, 1 ml, Pharmacia).

**SDS gel electrophoresis, staining and peptide sequencing.** Polyacrylamide gels (12) were silver-stained for glycoproteins (13). The purified active enzyme was fractionated by SDS-PAGE, stained with Coomassie Blue and the major band was excised for tryptic digestion,

<sup>1</sup> The sequence described in this paper was deposited with the GenBank as Accession No. AF044603. The work was supported by a grant from the Monsanto Corp., St. Louis, MO, USA and by the F21C program of the University of Missouri.

<sup>2</sup> Correspondence. Dr. R. O. Morris, Biochemistry Department, 117 Schweitzer Hall, University of Missouri, Columbia, Missouri 65211 USA. Fax: (573) 884-4695. E-mail: [morrisr@missouri.edu](mailto:morrisr@missouri.edu).

Abbreviations used: DCPIP, dichlorophenolindophenol; PCR, polymerase chain reaction; PVPP, polyvinylpyrrolidone.



HPLC and sequencing by Monsanto Co. or by the Biotechnology Resource Laboratory of Yale University.

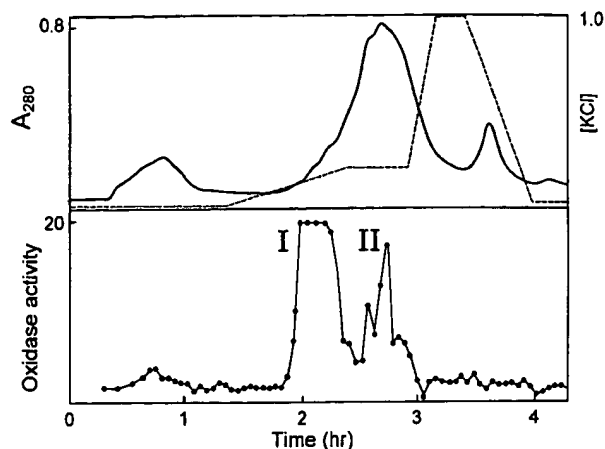
**Gene isolation.** Based on the sequence of selected tryptic peptides, degenerate inosine containing oligonucleotides were synthesized and used as primers. Primer C: TACGTIGAYGGIWSIGTITTCGT; primer F': TAIACRATIAGIGGCCICACRAT; primer D: ACIGAYCTIGCIAACACIGGITT; primer E': ATRTCIGTICCTGIA-GRATICG. Hot start touchdown PCR was carried out (14,15) with amplification for 40 cycles: annealing 65-45°C/20 sec (decreased by 1°C every other cycle), extension 72°C/30 sec and melting 94°C for 20 sec. Amplified fragments were ligated to pCRII DNA and transformed into *E. coli* INVαF' (Invitrogen, Carlsbad, CA, USA) and sequenced. The oxidase-specific PCR fragment CF400 was used to probe a maize genomic library.

A maize genomic library (partial *TaqI* genomic digest of Missouri 17 DNA in lambda-FIX II) was purchased from Stratagene (La Jolla, CA, USA). Phage were plated onto NZY agar plates and adsorbed (16) to sheets of Hybond N membrane (Amersham, Arlington Heights, IL, USA). The PCR product, CF400, was labeled with <sup>32</sup>P by primer extension using the Klenow fragment of DNA polymerase and primers C and F' and hybridized to membrane bound phage DNA (50% formamide, 5 × SSPE, 2 × Denhardt's solution, 0.2% SDS, 100-200 µg/ml denatured herring sperm DNA, 16 hours at 45°C). After three rounds of purification, one strongly positive phage, *λckx21*, was selected for characterization. Fragments were prepared by appropriate restriction enzyme digestion and subcloned into pBluescript II (Stratagene) for sequencing (Prism dideoxy terminator, Applied Biosystems, Foster City, CA, USA).

**Intron location.** After sequencing selected subclones of *λckx21*, the stochastic gene assembly program Exdomino (17) was used to find the best gene structures incorporating all strongly predicted exons and splice sites (splice site prediction was based on the SplicePredictor (18) program).

**Cross reaction of anti-CF400 antibody with maize oxidase.** Goat polyclonal antiserum was raised against the peptide expressed from the maize CF400 gene fragment in *E. coli*. After sodium sulfate purification (19), antibodies were coupled to Aminolink Plus gels (Pierce, Rockford, IL, USA). Activity depletion assays were performed by adding 0.2 ml (27 µg protein) of a crude oxidase preparation in phosphate-buffered saline to an immobilized antibody column, incubating at room temperature for 1 hour, eluting and assaying the eluate for oxidase activity.

**Expression of *ckx1* in *Pichia*.** Intron #2 (genomic sequence 3219-3312) was removed from *ckx1* using PCR splicing by overlap extension (20). Amplification#1, left primer, 5'-TGGGAATCCATGGGGAGATGGTGACGTGCTC-3'; right primer, 5'-GCCGTCCCACTGGATTGT-TGAGGGGTAGAC-3'. Amplification#2, left primer, 5'-CTCAACAAATCCATGTGGGACGACGGCATGTCCGGCG-3', right primer, 5'-GCCGTCTAGATCTAACTAAAACATGCATGGGCTATC-ATC-3'; final amplification, primers 5'-ATGGGAATCCATGGGGAGATGGTGACGTGCTC-3' and 5'-GCCGTCTAGATCTAACTAAAACATGCATGGGCTATC-3'. The final PCR product was cloned, sequenced and a *PflM1/Xba1* subfragment was substituted for the intron-containing *PflM1/Xba1* subfragment of the genomic clone. Intron #1 (genomic sequence 2113-2524) was then removed by replacement of the DNA between the restriction sites *PinA1* and *Nco1* with a linker constructed from the oligonucleotides 5'-CCGGTTTGGTACCGGT-3' and 5'-CATGACCGGTACCAAAA-3'. Extraneous linker-associated bases were removed by digestion with *PinA1* followed by re-ligation. The final intron-less fusion (containing the predicted signal peptide sequence) was ligated between the *PmlI/BsmBI* restriction sites of the *Pichia* cytoplasmic expression vector pPICZ-A (Invitrogen) transformed into *Pichia* strain X33 by electroporation, and selected on 100 µg/ml zeocin. One transformant (PPCKX1) was selected for expression studies. It was inoculated into BMGY (Invitrogen) medium (50 ml), grown overnight, resuspended in BMMY medium containing 0.5% (v/v) methanol, and grown at



**FIG. 1.** Chromatographic fractionation of cytokinin oxidase from maize kernels A: Bulk DEAE-cellulose chromatography of 1 Kg equivalent of the 45-65% ammonium sulfate fraction from immature maize kernels. Peak I was collected and pooled with material from several other experiments for further fractionation. Oxidase activity ( $\text{U } \mu\text{l}^{-1}$ ).

30°C with vigorous aeration. Additional methanol was added to 0.5% (v/v) at 24, 48 and 72 hours postinoculation. Samples were harvested for assay of cytokinin oxidase activity in cell lysates or in culture media.

**Physical studies.** Enzyme kinetic and spectral data were acquired on a HP 8453 diode array spectrometer and reduced using UV-Visible ChemStation software Rev A.05.02.

## RESULTS

### Purification of Maize Cytokinin Oxidase

The enzyme was extracted from a rich source: immature maize kernels, 5-8 days after pollination (21). It was purified by ammonium sulfate fractionation, DEAE-cellulose chromatography, concanavalin A-affinity chromatography, ion exchange chromatography on MonoQ and hydrophobic interaction chromatography. The presence of oxidase activity was detected up to and including the DEAE-cellulose chromatographic step by the method of Libreros-Minotta and Tipton (22). Subsequent purification stages were monitored either with this assay or with a rapid dichlorophenolindophenol (DCPIP) reduction assay (11). Figure 1 illustrates a representative separation at the DEAE stage.

Two peaks of oxidase activity were present (the major peak representing more than 70% of the total oxidase activity and a minor peak representing the balance). The major peak (Peak I) was selected for further characterization. It was retained completely on concanavalin A-agarose and was eluted by  $\alpha$ -methylmannoside. It is therefore glycosylated. Subsequent purification of this glycoprotein fraction on MonoQ and phenyl sepharose columns gave a highly purified oxidase preparation. Approximately 100 µg of purified oxidase was obtained from 40 Kg maize kernels. A

YVDGWVFNQ SLATDLANTG FFTDADVARI VALAGERNAT TVYSIEATLN YDNATAAAAA VDQELASVLG TLSYVEGFAP  
M-3  
QRDVAYTAFL DRVHGEEVAL NKLGLWRVPH PNLNMFVPRS RIADFDRGVF KGILQGTDIV GPLIVY  
M-4 M-1 Y-2

**FIG. 2.** Amino acid sequence of PCR product CF400. Double underlined regions locate tryptic peptides (M-3 and Y-2) used to design nested PCR primers. Single underlined regions indicate locations of other tryptic peptides (M-1 and M-4) derived from cytokinin oxidase confirming that CF400 was indeed oxidase-derived. The single residue difference between the observed and predicted sequences of M-4 was due to a single base change PCR error.

native gel of this preparation stained for enzyme activity (data not shown) gave a single active band that, when excised, subjected to SDS-PAGE and stained by the glycoprotein silver-staining method of Møller and Poulsen (13), gave a major band of approximately 60 kDa and several smaller minor bands. The 60 kDa band was selected for tryptic digestion and sequence analysis.

#### Isolation of an Oxidase Gene Fragment

Eight tryptic peptides were isolated from the purified oxidase and sequenced by Edman degradation. Of the eight, four proved useful: M-1, LGLWR; M-3, YVEGSVFXQSLATDLANTGFFT; M-4, VAYAAFLDR and Y-2, GILQGTDIVGPLIVY. Based on the sequences of M-3 and Y-2, inosine-containing degenerate oligonucleotide primers C, F', D and E' (Materials and Methods) were synthesized and used for nested touch-down PCR (14) with maize genomic DNA as template. When primers C and F' were employed, a ~440 bp DNA fragment (designated CF400) was produced. Re-amplification of CF400 with primers D and E' (predicted internal to C and F') gave a smaller (~400 bp) DNA fragment, suggesting that CF400 encompassed M-3 and Y-2. On cloning and sequencing CF400, it proved to encode not only M-3 and Y-2 at its proximal and distal ends but also peptides M-4 and M-1 as illustrated in Fig. 2.

#### Isolation and Characterization of the Oxidase Gene

The cloned PCR fragment, pCF400 was used to screen a Missouri 17 maize genomic library. One

MAVVYYLLLAGLIACSHALAGTPALGDDRGPWPASLAALALDGKLRDSTAAASTD  
FGNITISALPAAVLYPSSGTDLVALLSAANSTPGWPTYIAFRGRHSLMGOAFAPGGVVVN  
MASLGDAAPPRINVSADGRYVDAGGEQVWIDVLRASLARGVAPRSWNDYLYLTVGGTLS  
NAGISGOAFRHGPQISNVLEMDVITGHGEMVTCQKLNADLFDAVLGGLGQFGVITRARI  
AVEPAPARARWVRVYTDFAAFSADQERLTAPRPGGGASFGPMSYVEGSVFNQSLATD  
LANTGFFTDADVARI VALAGERNAT TVYSIEATLN YDNATAAAAAVDQELASVLGTLSSV  
EGFAQRDVAYAAFLDRVHGEEVALNKLGLWRVPH PNLNMFVPRSRIADFDRGVFKGILQ  
GTDIVGPLIVYPLNKSMDGMSAATPSEDVFYAVSLLFSSVAPNDLARLQEQNRRLRF  
CDLAGIQYKTYLARHTDRSDWRHFGAAKWNRFVEMKNKYDPKRLLSPGQDIFN

**FIG. 3.** Deduced amino acid sequence of *ckx1*. Double underline: signal peptide predicted by the method of Nielsen *et al.* (23); single underline: consensus FAD binding region (Prosite consensus PS 00862); shaded regions, consensus N-glycosylation sites.

strongly-hybridizing library phage,  $\lambda$ ckx21 was selected and subclones exhibiting homology to CF400 were further characterized and sequenced (GenBank Accession No. AF04460, 6733 bp). The programs Ex-domino and SplicePredictor (17,18) predicted the presence of an open reading frame encompassing CF400. The complete gene was designated *ckx1*. Three exons were present within *ckx1* at bases 1497-2112, 2524-3219 and 3321-3607 with a probable coding region start at nucleotide 1497 and with four in-frame stop codons beginning at 3605. Exon #2 contained CF400.

#### Predicted Oxidase Properties

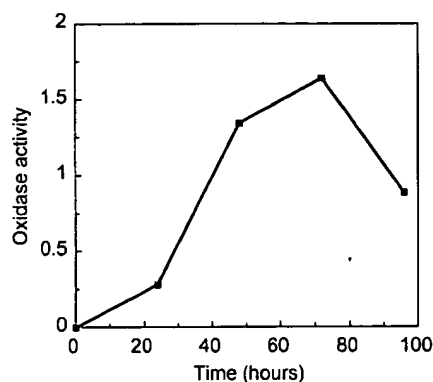
Figure 3 illustrates the predicted amino acid sequence of *ckx1*. The gene encodes a 57.2 kDa protein with a predicted eighteen amino acid signal peptide (23) and eight putative N-glycosylation sites. A BLAST comparison (24) with the GenBank, SwissProt, and PIR databases showed substantial homology with a number of flavoprotein oxygen oxidoreductases within the Prosite PS00862 consensus group, including 6-hydroxy-D-nicotine oxidase from *Arthrobacter oxidans* (25), and rat L-gulonolactone oxidase (26). The underlined sequence indicates the most likely FAD-binding domain. Laser desorption time-of-flight mass-spectrometric determination of the molecular weight of the expressed oxidase (see below) gave a provisional value of 69.620 kDa indicating the presence of approximately 10 hexose residues per consensus glycosylation site.

#### Proof that the *ckx1* Encodes Cytokinin Oxidase

Because the oxidase as purified from maize kernels was only 80% homogeneous by gel electrophoresis, it was important to demonstrate that it was encoded by *ckx1*. First, an antibody was raised against the protein

**TABLE 1**  
Recovery of Maize Cytokinin Oxidase from Immobilized Anti-CF400 Antibody Columns

Column	Percent activity recovered	
	Experiment 1	Experiment 2
Preimmune	64.5	60.3
Anti-CF400	3.2	2.7



**FIG. 4.** Expression of cytokinin oxidase in *Pichia*. An intronless construct was derived from *ckx1* and inserted into the *Pichia* cytoplasmic expression vector pPICZ-A under the control of the methanol-inducible promoter AOX. Methanol was added at zero time and cytokinin oxidase activity was assayed in the supernatant. No activity was detected in transformed *Pichia* cell extracts or in untransformed *Pichia* supernatants. Oxidase activity (U L<sup>-1</sup>).

fragment expressed by CF400 in *E. coli*, immobilized on a chromatographic matrix and found to selectively remove oxidase activity from a crude maize preparation (Table 1). More than 96% of added oxidase activity was retained by an anti-CF400 antibody column, whereas more than 60% of the activity passed through a control column of immobilized preimmune goat serum.

Second, when *ckx1* was expressed in *Pichia* (27), active cytokinin oxidase was secreted into the culture medium. To prepare the expression construct, introns were first removed by splice overlap extension (20). When the resulting intronless coding region was inserted into a *Pichia* cytoplasmic expression vector under the control of the alcohol oxygenase (AOX) promoter (28), large amounts of cytokinin oxidase activity

were detected in the culture media (Fig. 4). No oxidase activity was detected in induced *Pichia* cell homogenates, suggesting that the native plant signal peptide was being recognized and cleaved correctly by *Pichia*. Oxidase activity was not detected in control *Pichia* strains containing no insert nor in strains secreting human serum albumin.

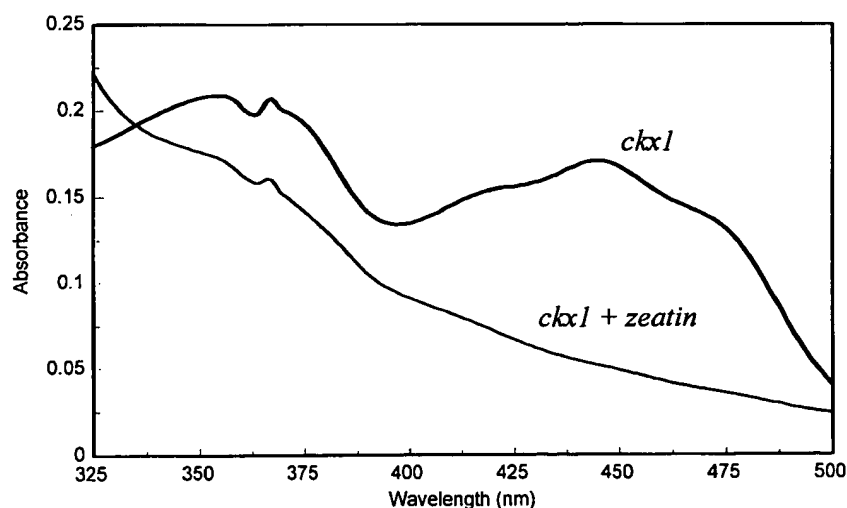
Preliminary kinetic data indicated that for the recombinant enzyme, the  $K_M$  for *trans*-zeatin was 19  $\mu$ M. As expected, *iso*-pentenyladenine and *iso*-pentenyladenosine were also substrates but dihydrozeatin was not. Thus, *ckx1* encodes a glycosylated cytokinin oxidase with enzymatic properties close to those reported for the native enzyme (29). Finally, the recombinant oxidase had a spectrum (Fig. 5) characteristic of a flavoprotein oxidoreductase (30). Absorbance maxima were present at 366 and 450 nm, and were bleached by addition of zeatin. Atomic absorption analysis for the presence of copper was negative.

#### Homology to Other Genes

*Ckx1* displays homology to two other genes (Fig. 6). The first is represented in a recently deposited *Arabidopsis* BAC clone (9) and is possibly the dicot homolog of the maize gene. There is more than 40% identity in the derived amino acid sequences between the two genes. More surprisingly, there is also extensive homology at the peptide level (33.6%) with the *fas5* open reading frame from *Rhodococcus fascians* (10).

#### DISCUSSION

The *ckx1* gene encodes a glycosylated FAD-containing oxidoreductase with substrate specificity closely resembling that reported for native maize cyto-



**FIG. 5.** UV absorbance spectrum of recombinant cytokinin oxidase. Recombinant cytokinin oxidase was expressed in *Pichia*, concentrated from the culture medium and purified by size exclusion chromatography. The absorption spectrum between 325 and 500 nm is shown. Upper trace: purified oxidase; lower trace: purified oxidase plus 5  $\mu$ M zeatin.

```

arab      : MGLTSSLRFHRRONKTFGLIFVILVLSCLPGRTNLCNHSVSTPKELPSSNPDSIRSSIVSDLEEVISDDVENVV--KDFGRYQLPLATIHPRSPDSSMMK
ckx1      : -----VAVVYYL LAGLIAASHALAAAGTPAGDDRGRPWPAASIAAALDKERTLSNATAAASDFGNITSLEAAVLYSSSTGDLVALLS
fas5      : -----MSIWHTDVHLTSAGDFGNCIHAKEPVVVVVERTVADVQEALE

arab      : HIVHLGSTSNI TVAARGHSHSLQSOALAHOGVVIMESI---RSDIRIYKGLPVVLVSFGLWINILPETIKYELSPKSWTDYHLLTVGGTLSNAGISGOAFKHG
ckx1      : -AANSIPGWPFYIAFRGRGHSIMSQPFFPGCVVVMASHGDAAAEPRIIVSADGRYVDAGGEQVWIDVLRASLARGVAPRSPNDYLYLTVGGTLSNAGISGOAFRHG
fas5      : Y---AAARNISLAVRGSCHSTYGCQCDGCVVLDMKRENTVH---DVFSGATIDAGVR--ASDVVAATISQOQTHPVLTIDYIGTVGGTLSVGFSSHGFE

arab      : PQINNVYQLFIVTGKGEVVTCSSEKRNSELEFSVLGGLGQGLITRARISLEPAHMMVKIRVLYSDSAFSRDQHYLSKEK-----TFDYVEGFYINR---H
ckx1      : PQISNVLEMDVITGHGEMVTCCKQLNADLEDAVLGGLGQFCVITRARIAVEPAARARWVRFVYTDHAAFSADQERTAPPGGGGASFGPMSVVEGSEVNVQSLAT
fas5      : LATTNVDSLAVVTGSDFRESAVSNSELEDAVRGGLGQFGVIVNETIRLTAHESVFOYKCYSNLGVLGQGLPANSN-----LTFHVQGRIRVFE

arab      : DILLANWRSSFSFNDSTQASRFKSUGKFLYCLEVVKYFH--PEEASSMDQETGKLLSEINVTPTSTLESSEVETIEFLDRVHLAERKIPAKGLAEVPPWLNLLIPKSS
ckx1      : DLAATGFFEDADVARIVLAGERNATTVYSLEATLNYNATAAAAVDOFLASVLGTLISYVEGFAPORDVATAAFLLDRVHGEVAINKLGLRVPHPWLNMEVPRSH
fas5      : -----ANCHLRRLDIKYETPPRR-----PDDDAILLSLQDSCAEINSDVDYGDFTNRADDEIDAEHTSEAFYPPHWSLLTHAD

arab      : IYDFATEVFNNTLSNN--NGPILYIPVNOCKKKKHTSHITINEDIEYLVAFLEAVFNSSGKNDIEYLLKORQVNFCAANINVEQYLPHYETQKEKSHFG--KR
ckx1      : IADFDRGVFKGTHQGTIVGPIIVYPLNKSMDDGMGAATISDDVFYAVSLLEFSS-----PVDIARIQENRRTIRFCDLAGIQYTYLRRHTDRSDVVRHECAAF
fas5      : IEDTETTSSTTDGLGNSILIMVYPIETP-ITAPFPIPHCTFFMLAVRIASEGAEE-----RMIASNRLIYEQARDVGVAYAVNVPMSPGDWCTHFG--SR

arab      : KFTFACRQAYDHLAILAPGORTKTTGKLSPIQLAKSKATGSPQRYHYASILPKPRTV
ckx1      : KNRFEVFNKYDEKRLSPGCDLHN-----
fas5      : KCAIRARRRFDHYRILAPCYRHSFD-----

```

FIG. 6. Homology between an *Arabidopsis* oxidoreductase, *ckx1*, and *fas5*. Peptide sequences were compared by the Pileup program (Genetics Computer Group, Madison, WI, USA) using the blosum62 comparison matrix, an open gap penalty of 12 and a gap extension penalty of 4. *Arab*: hypothetical oxidoreductase (bases 15517 to 17752) of the *Arabidopsis* chromosome II BAC clone T32G6 (9). *Ckx1*: the maize cytokinin oxidase; *Fas5*, hypothetical oxidoreductase from *Rhodococcus fascians* (10).

kinin oxidase (29,31). The presence of a consensus FAD binding domain of oxygen-dependent FAD oxidoreductases (Prosite class PS00862), suggested that *ckx1* may encode a flavoprotein oxidase and this was confirmed by the absorption spectrum of the recombinant enzyme that indicated the presence of substrate-reducible FAD. The enzyme has long been thought to be a copper-containing secondary amine oxidase (5). The data do not support this hypothesis. Although a topaquinone consensus sequence (NYD) (32) characteristic of copper-containing amine oxidases was noted for *ckx1*, this consensus was not present in the putative *Arabidopsis* homolog. Further, the enzyme contained no detectable copper by atomic absorption analysis. It remains to be determined whether there are other cofactors besides FAD or whether there are associated electron transfer proteins.

The availability of the *ckx1* gene will lead to an understanding of cytokinin oxidase in plant development. It will be of particular interest to determine *ckx1* expression patterns, not only within the developing maize kernel, but within other regions of the plant. The fact that the enzyme is glycosylated suggests that it is extracellular. With the availability of anti-CKX1 antibodies, it will now be possible to determine whether the enzyme is secreted to the apoplast or to protein bodies.

Perhaps the most interesting and perplexing finding relates to the homology between *ckx1* and *fas5* of *Rhodococcus fascians*. *R. fascians* causes shoot hypertrophy (Witches broom disease) in dicotyledonous plants. The bacteria possess a number of genes needed for pathogenicity (10). These include an *ipt* homolog and *fas5*. If indeed *fas5* does encode cytokinin oxidase activity, it is not evident how such activity should contribute to the pathological function of *Rhodococcus*, that has, up to this time, been believed to be due primarily to expression of the *ipt* gene resulting in excess production of cytokinin in infected tissues.

The isolation of *ckx1* represents the first characterization of a plant gene strictly involved in cytokinin metabolism. The only other cytokinin-metabolizing enzyme that has been fully identified, is an adenine phosphoribosyl transferase (APRT) that has been isolated from *Arabidopsis thaliana* (33). In this case, however, APRT substrate specificity includes bases other than cytokinins. APRT cannot only catalyze phosphoribosylation of free cytokinin bases such as zeatin, it can also utilize adenine as substrate. There is therefore some question as to whether it is strictly involved in cytokinin metabolism.

After this work was complete, similar findings were reported by Houba-Herlin, Pethe, D'Alayer, and Laloue (GenBank Accession # 3882018) and will be published shortly.

## ACKNOWLEDGMENTS

We thank Paige Patterson and Jean Cole for skilled technical assistance; Dale Blevins and Tim Reinbott for provision of plant material; Ming Hao, Timothy Reich, and Janice Edwards for provision of anti-CF400 antibody; Christine Smith for peptide sequence determination; Lila Barofsky for mass spectrometric analysis and Volker Brendel for access to the EXDOMINO program prior to publication.

## REFERENCES

1. Beaty, J. S., Powell, G. K., Lica, L., Regier, D. A., MacDonald, E. M. S., Hommes, N. G., and Morris, R. O. (1986) *Mol. Gen. Genet.* **203**, 274–280.
2. Barry, G. F., Rogers, S. G., Fraley, R. T., and Brand, L. (1984) *Proc. Natl. Acad. Sci., U. S. A.* **81**, 4776–4780.
3. Lichter, A., Manulis, S., Sagee, O., Gafni, Y., Gray, J., Meilan, R., Morris, R. O., and Barash, I. (1995) *Mol. Plant Microbe Interact.* **8**, 114–121.
4. Armstrong, D. J. (1994) in *Cytokinins. Chemistry, Action and Function* (Mok, D. W. S., and Mok, M. C., Eds), pp. 139–154, CRC Press, Boca Raton, Fla.
5. Hare, P. D., and Van Staden, J. (1994) *Physiol. Plant.* **91**, 128–136.
6. Kaminek, M., and Armstrong, D. J. (1990) *Plant Physiol.* **93**, 1530–1538.
7. Burch, L. R., and Horgan, R. (1989) *Phytochem.* **28**, 1313–1319.
8. Laloue, M., and Fox, J. E. (1989) *Plant Physiol.* **90**, 899–906.
9. Rounsley, S. D., Lin, X., Ketchum, K. A., Crosby, M. L., Brandon, R. C., Sykes, S. M., Kaul, S., Mason, T. M., et al. (1997) GenBank AC002510, bases 15517–17752.
10. Crespi, M., Vereecke, D., Temmerman, W., Van Montagu, M., and Desomer, J. (1994) *J. Bacteriol.* **176**, 2492–2501.
11. Morris, R. O., and Laskey, J. G., Unpublished.
12. Laemmli, U. K. (1970) *Nature* **227**, 680–685.
13. Moller, H. J., and Poulsen, J. H. (1995) *Anal. Biochem.* **226**, 371–374.
14. Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., and Mattick, J. S. (1991) *Nucl. Acid. Res.* **19**, 4008.
15. Mullis, K. B. (1991) *PCR Methods and Applications* **1**, 1–4.
16. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1990) *Molecular Cloning. A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press.
17. Brendel, V., and Kleffe, J. (1998) <http://gnomic.stanford.edu/~volker/SplicePredictor.html>.
18. Brendel, V., Kleffe, J., Carle-Uriste, J., and Walbot, V. (1998) *J. Mol. Biol.* **276**, 85–104.
19. Deutsch, H. (1967) in *Methods in Immunology and Immunohistochemistry I* (Williams, C. A., and Chase, M. W., Eds), pp. 315–321, Academic Press, San Diego.
20. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989) *Gene* **77**.
21. Dietrich, J. T., Kaminek, M., Blevins, D. G., Reinbott, T. M., and Morris, R. O. (1995) *Plant Physiol. Biochem.* **33**, 327–336.
22. Libreros-Minotta, C. A., and Tipton, P. A. (1995) *Anal. Biochem.* **231**, 339–341.
23. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Engineering* **10**, 1–6.
24. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
25. Brandsch, R., Hinkkanen, A. E., Mauch, L., Nagursky, H., and Decker, K. (1987) *Eur. J. Biochem.* **167**, 315–320.
26. Koshizaka, T., Nishikimi, M., Ozawa, T., and Yagi, K. (1988) *J. Biol. Chem.* **263**, 1619–1621.
27. Sreekrishna, K., Potenz, R. H., Cruze, J. A., McCombie, W. R., Parker, K. A., Nelles, L., Mazzaferro, P. K., Holden, K. A., et al. (1988) *J. Basic Micro.* **28**, 265–278.
28. Tschoop, J. F., Brust, P. F., Cregg, J. M., Stillman, C. A., and Gingeras, T. R. (1987) *Nucl. Acid. Res.* **15**, 3859–3876.
29. McGaw, B. A., and Horgan, R. (1983) *Planta* **159**, 30–37.
30. Bruhmuller, M., Mohler, H., and Decker, K. (1972) *Eur. J. Biochem.* **29**, 143–151.
31. Whitty, C. D., and Hall, R. H. (1974) *Can. J. Biochem.* **52**, 789–799.
32. Janes, S. M., Palcic, M. M., Scaman, C. H., Smith, A. J., Brown, D. E., Dooley, D. M., Mure, M., and Klinman, J. P. (1992) *Biochem.* **31**, 12147–12154.
33. Lee, D., and Moffatt, B. A. (1993) *Physiol. Plant.* **87**, 483–492.

- nitric oxide, cyclic GMP, and cyclic ADP-ribose, *Proc. Natl. Acad. Sci. U. S. A.* 95, 10328–10333
- 6 Delledonne, M. *et al.* (1998) Nitric oxide functions as a signal in plant disease resistance, *Nature* 394, 585–588
  - 7 Van Camp, W., Inzé, D. and Van Montagu, M. (1998) H<sub>2</sub>O<sub>2</sub> and NO: redox signals in disease resistance, *Trends Plant Sci.* 3, 330–334
  - 8 Bredt, D.S. *et al.* (1991) Cloned and expressed nitric oxide synthase structurally resembles cytochrome P-450 reductase, *Nature* 351, 714–719
  - 9 Yamasaki, H., Sakihama, Y. and Takahashi, S. (1999) An alternative pathway for nitric oxide production in plants: new features of an old enzyme, *Trends Plant Sci.* 4, 128–129
  - 10 Lipton, S.A. *et al.* (1993) A redox-based mechanism for the neuroprotective and neurodestructive effects of nitric oxide and related nitroso-compounds, *Nature* 364, 626–632
  - 11 Wink, D.A. *et al.* (1993) Nitric oxide protects against cellular damage and cytotoxicity from reactive oxygen species, *Proc. Natl. Acad. Sci. U. S. A.* 90, 9813–9817
  - 12 Dangel, J.F., Dietrich, R.A. and Richberg, M.H. (1996) Death don't have no mercy: cell death programmes in plant-microbe interactions, *Plant Cell* 8, 1793–1807

## Cytokinin oxidase strikes again

Research into the enzymes involved in several aspects of cytokinin metabolism (including storage and catabolic pathways) is undergoing an unprecedented period of success. We recently reported<sup>1</sup> on new findings concerning the biochemical and molecular properties of cytokinin oxidase (CKO), a poorly characterized, but crucial enzyme, which irreversibly inactivates cytokinins. Although CKO has been classified as a copper-containing amine oxidase<sup>1</sup>, clear evidence is now emerging that argues against this classification<sup>2–4</sup>. In particular, as previously described<sup>1</sup>, there is evidence that CKO from wheat is an FAD-containing flavoprotein (P. Galuszka *et al.*, unpublished), and recent advances have further extended this finding<sup>3,4</sup>.

The predominant CKO from immature maize kernels has now been purified and its structural gene (*ckx1*) has been isolated and characterized<sup>3</sup>. The gene encodes a glycosylated protein of ~57 kDa that possesses an 18 amino acid signal peptide and a consensus FAD-binding sequence<sup>3</sup>. By

expressing *ckx1* in *Pichia* it was possible to obtain an active glycosylated-CKO that was secreted into the culture medium. The absorption spectrum of the recombinant enzyme, which shows substrate reactivity similar to that reported for the native enzyme, indicated the presence of substrate-reducible FAD, whereas no copper ions were detected by atomic absorption analysis<sup>3</sup>. The search for homology to other genes revealed that *ckx1* has sequence homology (>40% identity in the derived amino acid sequences) with a putative oxidoreductase from *Arabidopsis thaliana*, and, to a lesser extent, with the *fas5* gene from *Rhodococcus fascians*<sup>5</sup>.

About the same time, the group led by Michel Laloue<sup>4</sup> reported similar results on maize CKO to those of Morris and colleagues<sup>3</sup>. In this study, photolabelling of maize CKO with a tritiated azido-derivative of *N*-(2-chloro-4-pyridyl)-*N'*-phenylurea (a potent cytokinin agonist and CKO inhibitor) facilitated purification of the enzyme by preparative 2D-gel electrophoresis. Subsequently, the CKO cDNA was isolated, cloned and expressed in moss (*Physcomitrella patens*) protoplasts to generate a catalytically active maize CKO (Ref. 4). The deduced CKO amino acid sequence shows sequence homology with an FAD-binding domain common to several oxidases. Moreover, a GHS motif was found in the maize CKO sequence, suggesting that the enzyme might covalently bind to FAD through a histidine residue<sup>4</sup>.

In conclusion, maize CKO has now been cloned and expressed in two independent laboratories using different approaches. This has confirmed that the enzyme is an FAD-dependent oxidase. Several pieces of evidence confirm, therefore, that the previous classification of CKO among copper amine oxidases can be dropped, and that CKO belongs instead to the large family of FAD amine oxidases.

The progress made on CKO comes at a time of significant advances in research on other enzymes involved in the control of the cytokinin cell cycle. The genes encoding two enzymes, *O*-glucosyltransferase<sup>5</sup> and *O*-xylosyltransferase<sup>6</sup>, from *Phaseolus lunatus* and *P. vulgaris*, respectively, which catalyse the formation of *O*-glycosyl derivatives of zeatin (the most active and ubiquitous of the naturally occurring cytokinins) have recently been isolated. The two genes exhibit 93% identity at the nucleotide level and the deduced amino acid sequence has 90% similarity<sup>5,6</sup>. Because glycosyl conjugates of zeatin are found in many plant tissues and are assumed to be important for transport and storage, and protection against degradative enzymes

(such as CKO), the cloning of these genes will benefit further studies on the regulation of cytokinin metabolism.

## References

- 1 Rinaldi, A.C. and Comandini, O. (1999) Cytokinin oxidase: new insight into enzyme properties, *Trends Plant Sci.* 4, 127–128
- 2 Galuszka, P. *et al.* (1998) Cytokinins as inhibitors of plant amine oxidase, *J. Enzyme Inhib.* 13, 457–463
- 3 Morris, R.O. *et al.* (1999) Isolation of a gene encoding a glycosylated cytokinin oxidase from maize, *Biochem. Biophys. Res. Commun.* 255, 328–333
- 4 Houba-Hérin, N. *et al.* (1999) Cytokinin oxidase from *Zea mays*: purification, cDNA cloning and expression in moss protoplasts, *Plant J.* 17, 615–626
- 5 Martin, R.C., Mok, M.C. and Mok, D.W.S. (1999) Isolation of a cytokinin gene, *ZOG1*, encoding zeatin *O*-glucosyltransferase from *Phaseolus lunatus*, *Proc. Natl. Acad. Sci. U. S. A.* 96, 284–289
- 6 Martin, R.C., Mok, M.C. and Mok, D.W.S. (1999) A gene encoding the cytokinin enzyme zeatin *O*-xylosyltransferase of *Phaseolus vulgaris*, *Plant Physiol.* 120, 553–558

Andrea C. Rinaldi\* and Ornella Comandini  
Dipartimento di Scienze Ambientali, Università dell'Aquila, Via Vetoio, Loc. Coppito, I-67100 L'Aquila, Italy

\*Author for correspondence  
(tel +39 0862 433209;  
fax +39 0862 433205;  
e-mail rinaldi@vaxca1.unica.it)

## Genetic nomenclature guidelines for the model legume *Lotus japonicus*

The legume *Lotus japonicus*, a diploid relative of Bird's-foot-trefoil (*L. corniculatus*) is widely used as a model legume for studying symbiotic interactions with *Mesorhizobium loti* and mycorrhizal fungi, such as *Glomus intraradices*<sup>1,2</sup>. Because the first papers describing the isolation and characterization of mutants in *L. japonicus* have either been published<sup>2–5</sup>, or are in preparation, it is time to establish a genetic nomenclature that provides consistency and clarity. To facilitate communication between research groups

**Extra example for increased seed yield:**

**Example A.**

*Nicotiana tabacum* L. cv. Samsun NN leaf explants were transformed with the vector Bin-Hyg-TX carrying the AtCKX1 gene or the AtCKX 2 gene under control of CaMV 35 S promoter. Several lines originating from these transformed plants were further cultivated and their seed size was analysed (Table 12).

Tobacco Plant having the transgene CKX1 and CKX2 all showed an increase in seed area, a parameter for seed size.

TABLE 12

Tobacco plant	Description	transgene	Average seed area
2	T1 38 nullizygote		0
3	T1 38 nullizygote		0.279
4	T1 38 nullizygote		0.297
5	WT		0.248
6	WT		0.243
7	WT		0.264
8	WT		0.277
1	T1 38 transgenic	CKX2	0.353
9	T1 38 transgenic	CKX2	0.281
10	T1 38 transgenic	CKX2	0.293
11	T1 38 transgenic	CKX2	0.329
12	T1 38 transgenic	CKX2	0.282
13	T1 8 transgenic	CKX1	0.278
14	T1 8 transgenic	CKX1	0.315
15	T1 8 transgenic	CKX1	0.322
16	T1 8 transgenic	CKX1	0.312

---

# Reaction Temperature Constraints in DNA Computing

---

Russell Deaton

The Department of Electrical Engineering  
The University of Memphis  
Memphis, TN, 38152  
rjdeaton@memphis.edu

## Abstract

Using the thermodynamics of DNA melting, a technique is proposed to choose a reaction temperature for the DNA computation that minimizes the potential for mishybridizations.

Adleman[Adleman, 1994] showed the computational potential of the hybridization reaction and other molecular biology protocols. A basic framework for a computation based on oligonucleotide template matching reactions, or hybridizations, consists of three steps: 1) encoding of the problem instance in DNA oligonucleotides such that solution is enabled with molecular biology protocols, 2) the basic processing of the pool of oligonucleotides with hybridization and ligation reactions, and 3) extraction of the result with separation techniques, such as polymerase chain reaction (PCR) or hybridization to probe sequences attached to magnetic beads.

Ideally, oligonucleotide hybridizations occur only between Watson-Crick complements. Depending upon the conditions under which the hybridization is done, however, base pairs that are not Watson-Crick complements, or mismatched base pair, can occur[Sambrook et al., 1989]. In addition, oligonucleotides can hybridize in various alignments that are shifted from the designed one. These mishybridizations can produce false positives, or solutions to the DNA computation that appear to be correct, but actually are not, and false negatives, or the failure to produce a solution to a problem with DNA computation when one actually exists. The effect of the reaction conditions is characterized as the hybridization stringency. In general, as the reaction temperature of the hybridization is increased up to a critical point, the stringency increases. The temperature at which half the population of perfectly matched

oligonucleotide hybrids will have dissociated into single strands is called the melting temperature,  $T_m$ . The melting temperature is determined from curves of UV absorbance versus temperature, and can be interpreted as the fraction of single strands versus temperature[Wetmur, 1997]. Under conditions of low stringency, oligonucleotides can hybridize with more mismatched base pairs and over shorter lengths than under conditions of high stringency. Therefore, under assumptions of perfect Watson-Crick hybridization or perfect Watson-Crick complementation between oligonucleotides, an effect of the hybridization stringency is to introduce a possibility of false positives and negatives through mismatched hybridizations and shifted hybridizations. These mishybridizations can occur either in the basic processing step (step 2 above) or in the extraction process (step 3 above).

In this paper, a method based on the thermodynamics of DNA melting is used to estimate the reaction temperature for a given oligonucleotide encoding of a problem. The nearest-neighbor base stacking model[Borer et al., 1974] for the melting temperature of short oligonucleotides is used. The estimated reaction temperature should minimize the potential for mishybridizations. An upper limit on the reaction temperature is estimated from the melting temperature and half-width of the melting curve of the oligonucleotide with the lowest melting temperature. A lower limit is set by requiring that the oligonucleotide with the highest melting temperature have minimum potential for mishybridizing with its nearest neighbor, as measured by the Hamming distance.

Based on a model of nearest neighbor stacking interactions, a formula for non-self complementary oligonucleotide melting temperature is given by[Borer et al., 1974],

$$T_m = \frac{\Delta H^\circ}{\Delta S^\circ - R \ln(C_t/4)}, \quad (1)$$



where  $\Delta H^\circ$  is the enthalpy change of the hybridization,  $\Delta S^\circ$  is the entropy change of the hybridization,  $R$  is the gas constant, and  $C_t$  is the concentration of the single-stranded oligonucleotides. Mismatched base pairs reduce the melting temperature, and therefore, the chance of mismatched hybridizations. In the classic estimate[Bonner et al., 1973],  $T_m$  decreases approximately 1°C per 1% mismatch between hybridizing oligonucleotides[Bonner et al., 1973].

The reaction temperature for a DNA computation must meet several criteria related to the melting temperatures of the oligonucleotides that represent the problem instance. To enable the computation, the reaction temperature should be chosen so that the pair of W-C complement oligonucleotides with the lowest melting temperature hybridize in sufficient number to allow their participation in the computation. Therefore, the reaction temperature should be less than the lowest melting temperature,  $T_m^l$ , plus 1/2 the width of the melting curve,  $\Delta T_{(T_m)}^l$ ,

$$T_r < T_m^l + \frac{1}{2} \Delta T_{(T_m)}^l. \quad (2)$$

By choosing the reaction temperature to allow hybridization of the oligonucleotide with the lowest melting temperature, the possibility of mishybridizations between oligonucleotides with higher melting temperatures is introduced. For example, suppose we choose  $T_r$  to allow hybridization of an  $n$ -mer composed of all A-T base pairs. Then, approximately, all consecutive G-C base pairs of length  $n/2$  can bind, producing potential frameshifted hybridizations. Likewise, hybridizations with mismatches are also possible. To minimize this potential for mishybridization,  $T_r$  should be chosen high enough to provide stringent conditions. To estimate the appropriate  $T_r$ , the rule of a 1°C decrease in the melting temperature for each 1% mismatch between the oligonucleotides is used. The percent mismatch between two oligonucleotides is the Hamming distance  $d$  between them under the operation of reverse Watson-Crick complementation divided by their length  $n$ [Deaton et al., 1998]. Therefore, the melting temperature of an oligonucleotide with another oligonucleotide, a Hamming distance  $d$  away is approximately

$$T_{mm} = T_m - 100 \frac{1^\circ\text{C}}{1\% \text{ mismatch}} \frac{d}{n}, \quad (3)$$

where  $T_{mm}$  is the melting temperature of the mismatched hybridization, and  $T_m$  is the melting temperature of perfect W-C complements. To minimize the potential for mismatched hybridizations, the reaction

temperature should be

$$T_r > T_{mm}^h + \frac{1}{2} \Delta T_{(T_{mm})}^h, \quad (4)$$

where the  $h$  superscript indicates the highest such melting temperature for a set of oligonucleotides.

Therefore, the reaction temperature should be chosen such that

$$T_{mm}^h + \frac{1}{2} \Delta T_{(T_{mm})}^h < T_r < T_m^l + \frac{1}{2} \Delta T_{(T_m)}^l. \quad (5)$$

This choice of reaction temperature should optimize the performance of a DNA computation with respect to mishybridizations involving mismatched base pairs.

In conclusion, the melting curves for oligonucleotide hybridization have been used to estimate a reaction temperature for a DNA computation which minimizes the potential for mishybridizations with mismatched base pairs. To summarize, the reaction temperature should be high enough to melt mishybridizations, and low enough to allow all planned hybridizations to occur in sufficient number to participate in the computation.

## References

- [Adleman, 1994] Adleman, L. M. (1994). Molecular computation of solutions to combinatorial problems. *Science*, 266:1021–1024.
- [Bonner et al., 1973] Bonner, T. I., Brenner, D. J., Neufeld, B. R., and Britten, R. J. (1973). Reduction in the rate of DNA reassociation by sequence divergence. *J. Mol. Biol.*, 81:123–135.
- [Borer et al., 1974] Borer, P. N., Dengler, B., Tinoco, Jr., I., and Uhlenbeck, O. C. (1974). Stability of ribonucleic acid double-stranded helices. *J. Mol. Biol.*, 86:843–853.
- [Deaton et al., 1998] Deaton, R., Garzon, M., Rose, J. A., Franceschetti, D. R., Murphy, R. C., and Jr., S. E. S. (1998). Reliability and efficiency of a DNA based computation. *Phys. Rev. Lett.*, 80:417–420.
- [Sambrook et al., 1989] Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, second edition.
- [Wetmur, 1997] Wetmur, J. G. (1997). Physical chemistry of nucleic acid hybridization. In *Preliminary Proceedings of the 3rd DIMACS Workshop on DNA Based Computers*, pages 1–25, Providence, RI. DIMACS, American Mathematical Society. Philadelphia, PA., June 23–27.

# Nonradioactive In Situ Hybridization Application Manual

*Roche Applied Science*

To select a chapter section, click on the corresponding page number.  
The chapters below are in Acrobat PDF format .

<b>Chapter 1</b> General Introduction to In Situ Hybridization	<b>General Introduction to <i>In Situ</i> Hybridization</b>	<u>9</u>
	Introduction to Hapten Labeling and Detection of Nucleic Acids	<u>11</u>
	Choosing the Right Labeling Method for your Hybridization Experiment	<u>15</u>

<b>Chapter 2</b> Guidelines for In Situ Hybridization	<b>Guidelines for <i>In Situ</i> Hybridization</b>	<u>21</u>
	Details of the Technique	<u>22</u>

<b>Chapter 3</b> Nucleic Acid Hybridization - General Aspects	<b>Nucleic Acid Hybridization - General Aspects</b>	<u>33</u>
--	---	-----------

<b>Chapter 4</b> Procedures for Labeling DNA, RNA, and Oligonucleotides with DIG, Biotin, or Fluorochromes	<b>Procedures for Labeling DNA, RNA, and Oligonucleotides with DIG, Biotin, or Fluorochromes</b>	<u>41</u>
	Random primed labeling of ds DNA with DIG-, Biotin- or Fluorescein-High Prime reaction mix	<u>41</u>
	PCR labeling of ds DNA with the PCR DIG Probe Synthesis Kit or PCR Labeling Mixes	<u>43</u>
	Nick-translation labeling of ds DNA with Nick Translation Mixes for <i>in situ</i> Probes	<u>49</u>
	Nick-translation labeling of ds DNA with DIG-, Biotin-, or Fluorochrome-labeled dUTP	<u>52</u>
	RNA labeling by <i>in vitro</i> transcription of DNA with DIG, Biotin or Fluorescein RNA Labeling Mix	<u>54</u>
	Oligonucleotide 3'-end labeling with DIG-ddUTP, Biotin-ddUTP, or Fluorescein-ddUTP	<u>58</u>
	Oligonucleotide tailing with a DIG-dUTP, Biotin-dUTP, or Fluorescein-dUTP mixture	<u>60</u>
	Chem-Link labeling of DNA or RNA with DIG or Biotin	<u>63</u>
	Estimating the Yield of DIG-labeled Nucleic Acids	<u>65</u>
	Purification of labeled probes using the High Pure PCR Product Purification Kit	<u>71</u>

<b>Chapter 5</b> Procedures for In Situ Hybridization to Chromosomes, Cells, and Tissue Sections	<b>Procedures for <i>In Situ</i> Hybridization to whole chromosomes</b>	<u>77</u>
	<i>In situ</i> hybridization to human metaphase chromosomes using DIG-, biotin-, or fluorochrome-labeled DNA probes and detection with fluorochrome conjugates	<u>78</u>
	Detection of chromosomal imbalances using DOP-PCR and comparative genomic hybridization (CGH)	<u>95</u>
	Fluorescence <i>in situ</i> hybridization of a repetitive DNA probe to human chromosomes in suspension	<u>104</u>

## Nucleic Acid Hybridization – General Aspects

This chapter discusses the effects of various components of the hybridization solution on the rate of renaturation and thermal stability of DNA hybrids free in solution. The features will be more or less identical to those of immobilized nucleic acids, such as in filter and *in situ* hybridizations. The largest deviation probably occurs in the kinetics. The reader is referred to the following literature for more background information: Casey and Davidson (1976), Cox et al. (1984), Flavell et al. (1974), Hames and Higgins (1985), Maniatis et al. (1982), Raap et al. (1986), Schildkraut and Lifson (1965), Spiegelman et al. (1973), Wetmur and Davidson (1968), Wetmur (1975).

### The main parameters that influence hybridization

Hybridization depends on the ability of denatured DNA to reanneal with complementary strands in an environment just below their melting point ( $T_m$ ).

The  $T_m$  is the temperature at which half the DNA is present in a single-stranded (denatured) form. The  $T_m$  value is different for genomic DNA isolated from various organisms, e.g., for *Pneumococcus* DNA it is 85°C, for *Serratia* DNA it is 94°C. The  $T_m$  can be calculated by measuring the absorption of ultraviolet light at 260 nm. The stability of the DNA is directly dependent on the GC content. The higher the molar ratio of GC pairs in a DNA, the higher the melting point.

$T_m$  and renaturation of DNA are primarily influenced by four parameters:

- ▷ Temperature
- ▷ pH
- ▷ Concentration of monovalent cations
- ▷ Presence of organic solvents

#### Temperature

The maximum rate of renaturation (hybridization) of DNA is at 25°C. However, the bell-shaped curve relating renaturation rate and temperature is broad, with a rather flat maximum from about 16°C to 32°C below  $T_m$ .

#### pH

From pH 5 - 9, the rate of renaturation is fairly independent of pH. Buffers containing 20 - 50 mM phosphate, pH 6.5 - 7.5 are frequently used.

*Note: Higher pH can be used to produce more stringent hybridization conditions.*

#### Monovalent cations

Monovalent cations (e.g., sodium ions) interact electrostatically with nucleic acids (mainly at the phosphate groups) so that the electrostatic repulsion between the two strands of the duplex decreases with increasing salt concentration, i.e. higher salt concentrations increase the stability of the hybrid. Low sodium concentrations affect the  $T_m$ , as well as the renaturation rate, drastically.

Sodium ion ( $\text{Na}^+$ ) concentrations above 0.4 M only slightly affect the rate of renaturation and the melting temperature.

The following equation has been given for the dependence of  $T_m$  on the GC content and the salt concentration (for salt concentrations from 0.01 to 0.20 M):

$$T_m = 16.6 \log M + 0.41 (\text{GC}) + 81.5$$

where M is the salt concentration (molar) and GC, the molar percentage of guanine plus cytosine. Above 0.4 M  $\text{Na}^+$ , the following formula holds:

$$T_m = 81.5 + 0.41 (\text{GC})$$

Free divalent cations strongly stabilize duplex DNA. Remove them from the hybridization mixture or complex them (e.g., with agents like citrate or EDTA).

## Nucleic Acid Hybridization – General Aspects

### Formamide

DNA melts (denatures) at 90° – 100°C in 0.1 – 0.2 M Na<sup>+</sup>. For *in situ* hybridization this implies that microscopic preparations must be hybridized at 65° – 75°C for prolonged periods. This may lead to deterioration of morphology. Fortunately, organic solvents reduce the thermal stability of double-stranded polynucleotides, so that hybridization can be performed at lower temperatures in the presence of formamide.

Formamide has for years been the organic solvent of choice. It reduces the melting temperature of DNA-DNA and DNA-RNA duplexes in a linear fashion by 0.72°C for each percent formamide. Thus, hybridization can be performed at 30° – 45°C with 50% formamide present in the hybridization mixture. The rate of renaturation decreases in the presence of formamide. The melting temperature of hybrids in the presence of formamide can be calculated according to the following equation:

For 0.01 – 0.2 M Na<sup>+</sup>:

$$T_m = 16.6 \log M + 0.41 (\text{GC}) + 81.5 - 0.72 (\% \text{ formamide})$$

For Na<sup>+</sup> concentrations above 0.4 M:

$$T_m = 81.5 + 0.41 (\text{GC}) - 0.72 (\% \text{ formamide})$$

To obtain a large increase of *in situ* hybridization signal for rDNA, hybridize with rRNA in 80% formamide at 50° – 55°C, instead of 70% formamide at 37°C.

Finally, it should be mentioned that during the *in situ* hybridization procedure, relatively large amounts of DNA can be lost (Raap et al., 1986).

## Additional hybridization variables

Additional parameters must be considered when calculating the optimal hybridization conditions including the probe length, probe concentration, the inclusion of dextran sulfate, the extent of mismatch between probe and target, the washing conditions, and whether the probes will be single- or double-stranded.

### Probe length

The rate of the renaturation of DNA in solution is proportional to the square root of the (single-stranded) fragment length. Consequently, maximal hybridization rates are obtained with long probes. However, short probes are required for *in situ* hybridization because the probe has to diffuse into the dense matrix of cells or chromosomes. The fragment length also influences thermal stability. The following formula, which relates the shortest fragment length in a duplex molecule to change in  $T_m$ , has been derived:

Change in  $T_m \cdot n = 500$  ( $n$  = nucleotides).

### Probe concentration

The probe concentration affects the rate at which the first few base pairs are formed (nucleation reaction). The adjacent base pairs are formed afterwards, provided they are in register (zippering). The nucleation reaction is the rate limiting step in hybridization. The kinetics of hybridization is considered to be a second order reaction [ $r = k_2$  (DNA) (DNA)]. Therefore, the higher the concentration of the probe, the higher the reannealing rate.

### Dextran sulfate

In aqueous solutions dextran sulfate is strongly hydrated. Thus, macromolecules have no access to the hydrating water, which causes an apparent increase in probe concentration and consequently higher hybridization rates.

### Base mismatch

Mismatching of base pairs results in reduction of both hybridization rates and thermal stability of the resulting duplexes. To discriminate maximally between closely related DNA sequences, hybridize under fairly stringent conditions (e.g. at  $T_m - 15^\circ\text{C}$ ). On the average, the  $T_m$  decreases about  $1^\circ\text{C}$  per % (base mismatch) for large probes. Mismatching in oligonucleotides greatly influences hybrid stability; this forms the basis of point mutation detection.

### Stringency washes

During hybridization, duplexes form between perfectly matched sequences and between imperfectly matched sequences. The extent to which the latter occurs can be manipulated to some extent by varying the stringency of the hybridization reaction. (See above.)

To remove the background associated with nonspecific hybridization, wash the sample with a dilute solution of salt. The lower the salt concentration and the higher the wash temperature, the more stringent the wash.

In general, greater specificity is obtained when hybridization is performed at a high stringency and washing at similar or lower stringency, rather than hybridizing at low stringency and washing at high stringency.

### Use of single-stranded versus double-stranded probes

A number of competing reactions occur during *in situ* hybridization with double-stranded probes. These include:

- Probe renaturation in solution
- *In situ* hybridization
- *In situ* renaturation (possibly, for ds targets)

Consequently, the use of single-stranded probes has advantages for *in situ* hybridization. Such probes can be made by using the single-stranded M13 (or like bacteriophage cloning vectors) as template, or by using transcription vectors which permit the production of large amounts of single-stranded RNA. (See Chapter 4 for a detailed description).

### Competition *in situ* hybridization

Recombinant DNA isolated from eukaryotic DNA often contains genomic repetitive sequences (e.g., the Alu sequence in humans). *In situ* hybridization to chromosomes with a probe which contains repetitive DNA usually results in uniform staining. However, unlabeled competitor DNA (usually total genomic DNA) prevents the repetitive probe sequences from annealing to the target, and leads to stronger *in situ* hybridization signals from the unique sequences in the probe. (This approach was first described for *in situ* hybridization by Landegent et al., 1987; Lichter

et al., 1988a; Pinkel et al., 1988.) Obviously, the greater the complexity of probe (plasmids < phages < cosmids < yeast artificial chromosomes < chromosome libraries), the greater the need for competition *in situ* hybridization. This approach has proved particularly useful for *in situ* hybridization with DNA isolated from chromosome-specific libraries (CISS-hybridization); a specific chromosome can be fluorescently labeled over its full length (Lichter et al., 1988a,b; Cremer et al., 1988; Pinkel et al., 1988).

### Oligonucleotide hybridization

The rules given for hybrid stability and kinetics of hybridization can probably not be extrapolated to hybridization with oligodeoxynucleotides. For *in situ* hybridization, the advantages of oligonucleotides include their small size (good penetration properties) and their single-strandedness (to prevent probe reannealing, as outlined in Chapter 1).

The small size, however, is also a disadvantage because it covers less target. The nonradioactive label should be positioned at the 3' or the 5' end; internal labeling affects the  $T_m$  too much.

In an experiment with 20-mers of 40 – 60% GC content, start with the hybridization conditions described below. Depending on the results obtained, you may decide to use other stringency conditions.

### Standard *in situ* hybridization conditions

Department of Cytochemistry and Cytometry,  
University of Leiden, Netherlands.

#### For "large" DNA probes ( $\geq 100$ bp):

- 50% deionized formamide
- 2x SSC (see below)
- 50 mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$  buffer; pH 7.0
- 1 mM EDTA
- carrier DNA/RNA (1 mg/ml each)
- probe (approx. 20 – 200 ng/ml)

#### Optional components:

- 1x Denhardt's (see below)
- dextran sulfate, 5 – 10%
- Temperature: 37° – 42°C
- Hybridization time: 5 min – 16 h

#### For synthetic oligonucleotides:

- 25% formamide
- 4x SSC (see below)
- 50 mM  $\text{NaH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$  buffer; pH 7.0
- 1 mM EDTA
- carrier DNA/RNA (1 mg/ml each)
- probe (approx. 20 – 200 ng/ml)
- 5x Denhardt's (see below)
- Temperature: room temperature
- Hybridization time: 2 – 16 h

#### Composition of SSC and Denhardt's solution

1x SSC: 150 mM NaCl, 15 mM sodium citrate; pH 7.0:  
Make a 20x stock solution (3 M NaCl, 0.3 M sodium citrate).

50x Denhardt's:  
1% polyvinylchloride, 1% pyrrolidone,  
2% BSA.

## References

- Casey, J.; Davidson, N. (1976) Rates of formation and thermal stabilities of RNA-DNA and DNA-DNA duplexes at high concentrations of formamide. *Nucleic Acids Res.* **4**, 1539-1552.
- Cox, K. H.; DeLeon, D. V.; Angerer, L. M.; Angerer, R. C. (1984) Detection of mRNAs in sea urchin embryos by *in situ* hybridization using asymmetric RNA probes. *Develop. Biol.* **101**, 485-503.
- Cremer, T.; Lichter, P.; Borden, J.; Ward, D. C.; Manuelidis, L. (1988) Detection of chromosome aberrations in metaphase and interphase tumor cells by *in situ* hybridization using chromosome specific library probes. *Hum. Genet.* **80**, 235-246.
- Flavell, R. A.; Birfelder, J. E.; Sanders, J. P. M.; Borst, P. (1974) DNA-DNA hybridization on nitrocellulose filters. I. General considerations and non-ideal kinetics. *Eur. J. Biochem.* **47**, 535-543.
- Hames, B. D.; Higgins, S. J. (1985) *Nucleic Acid Hybridization: A Practical Approach*. Oxford: IRL Press.
- Landegent, J. E.; Jansen in de Wat; Dirks, R. W.; Baas, F.; van der Ploeg, M. (1987) Use of whole cosmid cloned genomic sequences for chromosomal localization by nonradioactive *in situ* hybridization. *Hum. Genet.* **77**, 366-370.
- Lichter, P.; Cremer, T.; Borden, J.; Manuelidis, L.; Ward, D. C. (1988a) Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* suppression hybridization using recombinant DNA libraries. *Hum. Genet.* **80**, 224-234.
- Lichter, P.; Cremer, T.; Tang, C. C.; Watkins, P. C.; Manuelidis, L.; Ward, D. C. (1988b) Rapid detection of human chromosomes 21 aberrations by *in situ* hybridization. *Proc. Natl. Acad. Sci. USA* **85**, 9664-9668.
- Maniatis, T.; Fritsch, E. F.; Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratories.
- Pinkel, D.; Landegent, J.; Collins, C.; Fuscoe, J.; Segreaves, R.; Lucas, J.; Gray, J. W. (1988) Fluorescence *in situ* hybridization with human chromosome specific libraries: detection of trisomy 21 and translocations of chromosome 4. *Proc. Natl. Acad. Sci. USA* **85**, 9138-9142.
- Raap, A. K.; Marijnen, J. G. J.; Vrolijk, J.; Van der Ploeg, M. (1988) Denaturation, renaturation, and loss of DNA during *in situ* hybridization procedures. *Cytometry* **7**, 235-242.
- Schildkraut, C.; Lifson, S. (1965) Dependence of the melting temperature of DNA on salt concentration. *Biopolymers* **3**, 195-208.
- Spiegelman, G. B.; Haber, J. E.; Halvorson, H. O. (1973) Kinetics of ribonucleic acid-deoxyribonucleic acid membrane filter hybridization. *Biochemistry* **12**, 1234-1242.
- Wetmur, J. G. (1975) Acceleration of DNA renaturation rates. *Biopolymers* **14**, 2517-2524.
- Wetmur, J. G.; Davidson, N. (1988) Kinetics of renaturation of DNA. *J. Mol. Biol.* **31**, 349-370.



## Sequence Revision History

PubMed

Nucleotide

Protein

Genome

Structure

PMC

Taxonomy

OMIM

Books

Find (Accessions, GI numbers or Fasta style SeqIds)

About Entrez

difference between I and II as



Entrez

### Revision history for AC005917

Search for Genes  
LocusLink provides curated  
information for human, fruit  
fly, mouse, rat, and  
zebrafish

Help|FAQ

Batch Entrez: Upload a  
file of GI or accession  
numbers to retrieve  
protein or nucleotide  
sequences

Check sequence  
revision history

How to create WWW  
links to Entrez

LinkOut

Cubby

Related resources

BLAST

Reference sequence project

LocusLink

Clusters of orthologous groups

Protein reviews on the web

GI	Version	Update Date	Status	I	II
20197478	3	<a href="#">Apr 18 2002 11:43 PM</a>	Live	<input checked="" type="radio"/>	<input type="radio"/>
6598497	2	<a href="#">Apr 10 2000 5:46 PM</a>	Dead	<input type="radio"/>	<input checked="" type="radio"/>
6598497	2	<a href="#">Jan 11 2000 7:03 PM</a>	Dead	<input type="radio"/>	<input type="radio"/>
6598497	2	<a href="#">Dec 23 1999 12:25 PM</a>	Dead	<input type="radio"/>	<input type="radio"/>
6598497	2	<a href="#">Dec 17 1999 6:00 PM</a>	Dead	<input type="radio"/>	<input type="radio"/>
4191771	1	<a href="#">Jan 28 1999 6:04 PM</a>	Dead	<input type="radio"/>	<input type="radio"/>
4191771	1	<a href="#">Jan 27 1999 6:04 PM</a>	Dead	<input type="radio"/>	<input type="radio"/>
3900812	n/a	<a href="#">Nov 21 1998 6:05 PM</a>	Dead	<input type="radio"/>	<input type="radio"/>
3893062	n/a	<a href="#">Nov 19 1998 6:11 PM</a>	Dead	<input type="radio"/>	<input type="radio"/>
3831436	n/a	<a href="#">Nov 3 1998 6:03 PM</a>	Dead	<input type="radio"/>	<input type="radio"/>

Accession AC005917.3 was first seen at NCBI on Nov 3 1998 6:03 PM

[Disclaimer](#) | [Write to the Help Desk](#)  
[NCBI](#) | [NLM](#) | [NIH](#)



LOCUS AC005917 92721 bp DNA linear PLN 05-APR-2000  
 DEFINITION Arabidopsis thaliana chromosome II section 113 of 255 of the  
 complete sequence. Sequence from clones F27F23, F3P11.  
 ACCESSION AC005917 AE002093  
 VERSION AC005917.2 GI:6598497  
 KEYWORDS HTG.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 92721)  
 AUTHORS Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D.,  
 Fujii,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E.,  
 Feldblyum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Ronning,C.M.,  
 Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., VanAken,S.E., Umayam,L.,  
 Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H.,  
 Goodman,H.M., Somerville,C.R., Copenhagen,G.P., Preuss,D.,  
 Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and  
 Venter,J.C.  
 TITLE Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana  
 JOURNAL Nature 402 (6763), 761-768 (1999).  
 MEDLINE 20083487  
 PUBMED 10617197  
 REFERENCE 2 (bases 1 to 92721)  
 AUTHORS Lin,X.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712  
 Medical Center Dr., Rockville, MD 20850, USA

COMMENT [WARNING] On Apr 18, 2002 this sequence was replaced by a newer  
 version gi:20197478.  
 On Dec 17, 1999 this sequence version replaced gi:4191771.  
 The sequence and annotation of chromosome 2 were merged from those  
 of the individual clones on this chromosome after removing  
 overlaps. For detailed information, please see the TIGR web site  
 (<http://www.tigr.org/tdb/at/at.html>).

Genes were identified by a combination of three methods: Gene  
 prediction programs including GRAIL  
 (<ftp://arthur.epm.ornl.gov/pub/xgrail>), Genefinder (Phil Green,  
 University of Washington), Genscan (Chris Burge,  
<http://gnomic.stanford.edu/GENSCANW.html>), and NetPlantGene  
 (<http://www.cbs.dtu.dk/services/NetGene2/>); searches of the  
 complete sequence against a peptide database and plant EST  
 databases at TIGR, and manual curations based on those analyses.  
 Annotated genes are named to indicate the level of evidence for  
 their annotation. Genes with similarity to other proteins are named  
 after the database hits. Genes without significant peptide  
 similarity but with EST similarity are named as 'unknown' proteins.  
 Genes without protein or EST similarity, that are predicted by two  
 or more gene prediction programs over most of their length are  
 annotated as 'hypothetical' proteins. Genes encoding tRNAs are  
 predicted by tRNAscan-SE (Sean Eddy,  
<http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were  
 identified by repeatmasker (Arian Smit,  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are  
 numbered from the top to bottom of the chromosome.

We thank the CSHL/WashU/ABI consortium for sequencing BAC clones  
 F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone  
 F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards  
 and Satoshi Tabata for helpful assistance. In addition, we would  
 like to thank the TIGR Bioinformatics Department, especially Lixin  
 Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy  
 Peterson, Michael Holmes, and Delwood Richardson for software and  
 database support.

This work was supported by the National Science Foundation,  
 Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.

```

FEATURES             Location/Qualifiers
    source             1..92721
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /cultivar="Columbia"
                        /db_xref="taxon:3702"
                        /chromosome="II"
    misc_feature       <1..1671
                        /note="Sequence from clone F27F23"
    gene               complement(<1311..>5422)
                        /gene="At2g19410"
                        /note="synonym: F27F23.21; contains a protein kinase
                        domain profile (PDOC00100)"
    mRNA               complement(join(<1311..1511,1608..2357,2445..2873,
                        3102..3203,3293..3567,3957..4070,4242..4380,4652..4824,
                        5200..>5422))
                        /gene="At2g19410"
    CDS                complement(join(1311..1511,1608..2357,2445..2873,
                        3102..3203,3293..3567,3957..4070,4242..4380,4652..4824,
                        5200..5422))
                        /gene="At2g19410"
                        /codon_start=1
                        /product="putative protein kinase"
                        /protein_id="AAD10166.2"
                        /db_xref="GI:6598498"
                        /translation="MVVMLTQEMSGGGPKAEEGQLFVAVAVKGLIGDKLGGAGSRRRA
                        VRWAVDNLLPKADKFVMIHVITITSIPTPNILILMFTRMWVVTAGDRLPVEEVEESV
                        VEMYVRDVKKEYETVFVPFLKMCKSTRSTKRYFRSRRTKGTGVPLTVLRYAPETCEVY
                        IVCKDRITTKSMDPLINREPTSPHAAATAHDFLRDWAASFHTLRSPTLPDPRQSTEA
                        GTRRSASARELRFEALSLTCNPKPTPQSSKASSATTPEIFRRRRGSDIPQLNYSDFDK
                        TCTKPQSNVENIVSEHRDSDRSPPETSRKSKKVEIEEEVERLKNELOSTVFKYKQACE
                        ELFSTQNKVKMLSTEYLNESKRVNNAVEKEELQRNTAALEKERYMKAVKEVETAKALL
                        AREFCQRQIAEVNALRTYLEKKKVIDQLLGTDRHYRKYTIEEIVTATEGFSPEKVIGE
                        GGYGKVYQCSLDSTPAAVKVVRDLTPPEKKQEFLEVEVLSQLRHPHVLLLGACPENG
                        CLVYEYLENGSLEEYIFHRKNKPPLPWFIRFRVIFEVACGLAFLHSSKPEPIVHRDLK
                        PGNILLNRNVYSKIADVGLAKLVTVDVAPDNVTMYRNSVLAGTLHYIDPEYHRTGTIRP
                        KSDLYAFGIILQLLTARNPSGIVPAVENAVKKGTLTEMLDKSVTDWPLAETEELARI
                        GLKCAEFRCDRDPDLKSEVIPVLKRLVETANSKVKKEGSNLRAPSHYFCPIPREIMEE
                        PEIAADGFTYERKAILAWLEKHNISPVTRQKLDHFKLTPNHTLRSAIRDWKSRVRF SN
                        VVV"
    misc_feature       1672..>92721
                        /note="Sequence from clone F3P11"
    gene               <8111..>10293
                        /gene="At2g19420"
                        /note="synonym: F3P11.2; predicted by genscan"
    mRNA               join(<8111..8129,8632..8742,9989..>10293)
                        /gene="At2g19420"
    CDS                join(8111..8129,8632..8742,9989..10293)
                        /gene="At2g19420"
                        /note="hypothetical protein"
                        /codon_start=1
                        /protein_id="AAD10141.1"
                        /db_xref="GI:4191772"
                        /translation="MHEKVPKKTYDSELYNGYLAIYNGYGDLRYFTVKIENDICVSLV
                        WNTGMTKKSRGKKRIRINVYRPKIAPAKNKKPKPTKEQLMDPEFSDEDVLTSLGFDDG
                        GYIIPLPKKSMLINLPDSFRDELQTMHVSFYLRVGVKLMFG"
    repeat_region      8188..8269
                        /rpt_family="POLY_A"
    repeat_region      8824..8918
                        /rpt_family="(TA)n"
    gene               <11627..>14150
                        /gene="At2g19430"
                        /note="synonym: F3P11.3; contains a Trp-Asp (WD-40) repeat
                        signature"
    mRNA               join(<11627..11952,12040..12116,12198..12243,12371..12431,
                        12520..12573,12664..12744,12836..12875,13148..13248,
                        13421..13531,14061..>14150)
                        /gene="At2g19430"
    CDS                join(11627..11952,12040..12116,12198..12243,12371..12431,
                        12520..12573,12664..12744,12836..12875,13148..13248,

```

```

13421..13531,14061..14150)
/gene="At2g19430"
/codon_start=1
/product="putative WD-40 repeat protein"
/protein_id="AAD10142.1"
/db_xref="GI:4191773"
/translation="MYGDATNWNEDEYRESILKEREIETRTVFR TAWAPPARISNPDA
FVVASSDGT LAFHSLNSLVSQSASFGYSKGQDVMVAEPERVVRAHEGPAYDVKFYGED
EDALLSCGDDGRVGRGWKREFAESDVSLHLKENHLKPLEELINPQHKGPGWALS PMP
EINAMSVDPQSGSVFTAAGDSCAYCWDVESGKIKMTFKGHS DYLHTVVS RSSASQILT
GSEDGTARIWDCKTGKCVKVIQSQDKSRLRVSSMALDGS ESWLVCQGQKNLALWNL P
ASECVQTIPIPAHVQDVMFDEKQVAVGGYGGIVDVISQFGSHLCTFRSSSL"
gene      complement(<14574..>16216)
          /gene="At2g19440"
          /note="synonym: F3P11.4"
mRNA      complement(join(<14574..15187,15255..15754,15894..>16216))
          /gene="At2g19440"
CDS       complement(join(14574..15187,15255..15754,15894..16216))
          /gene="At2g19440"
          /codon_start=1
          /product="putative beta-1,3-glucanase"
          /protein_id="AAD10143.1"
          /db_xref="GI:4191774"
          /translation="MNLLAFVVGFGIMGIVMVDGLGVNWGMTATHKLPPKKVVQMLKD
NNINKVKLFDADET TMSALSGSGLEVMVAIPNDQLKVMGSYDRAKDWVHKVTRYNFN
GGVNITFVAVGNEPFLKSYNGSFINLTFPALQNIQNALNEAGLGSSVKATVPLNADV
DSPSSNPVPSAGRFRPDIIGQMTQIVDFLGNN SAPITINIYPFLSLYGNDDFPLNYAF
FDGAKPVDDNGIAYTNVFDANFDTLVSALKAVGHGDMPIIVGEVGPTEGDKHANS GS
AYRFYNGLLPRLGENRGTPLRPTYIEVYLFGLLDEDAKSIAPGEFERHWGIFKFDGQP
KFPIDLSGQGQNKLLIGAENVTYQPKKWCMFNTEAKDLTKLAANIDYACTFSDCTALG
YGSSCNTLDANGNASYAFNMYFQVKNQDEDACIFQGLATITTKNISQGCNFP IQIVA
STASSFSSSLVLLIAGVWFLLSGVMFEV"
repeat_region complement(17218..17335)
          /rpt_family="(TAAA)n"
repeat_region complement(17283..17395)
          /rpt_family="(TAAAA)n"
repeat_region 20638..20663
          /rpt_family="(TA)n"
repeat_region 21498..21547
          /rpt_family="(TA)n"
repeat_region 22410..22543
          /rpt_family="(TAAAA)n"
gene      <22846..>25865
          /gene="At2g19450"
          /note="synonym: F3P11.5"
mRNA      join(<22846..23235,23320..23400,23958..24113,24194..24271,
24350..24451,24552..24688,24833..24914,25019..25039,
25124..25186,25270..25317,25441..25503,25761..>25865)
          /gene="At2g19450"
CDS       join(22846..23235,23320..23400,23958..24113,24194..24271,
24350..24451,24552..24688,24833..24914,25019..25039,
25124..25186,25270..25317,25441..25503,25761..25865)
          /gene="At2g19450"
          /codon_start=1
          /product="diacylglycerol O-acyltransferase"
          /protein_id="AAD10144.1"
          /db_xref="GI:4191775"
          /translation="MAILDSAGVTTVTENG GGEFVDLDR LRRRKRSDSSNGLLLSGS
DNNSP SDDVGAPADVRDRIDS VVNDDAQGTANLAGDNNGGGDNNGGGRGGGEGRG NAD
ATFTYRPSVPAHRRARESPLSSDAIFKQSHAGLFNLCVVVLI AVNSRLIIENLMKLS D
PVILKMDRCD SAFLSGVTLMLLTCIVWLKLVSYAHTSYDIRSLANAADKANPEVSYYV
SLKSLAYFMVAPTLCYQPSYPR SACIRKGWVARQFAKLVI FTGFMGFIIIEQYINPIVR
NSKHPLKGDLLYAIERVLKLSVPNLVYVWLCMFYCF FHLWLNLIAELLCFGDREFYKDW
WNAKSVGDYWRMWNMPVHKWMVRHIYF PCLRSKIPKTLAII IAFLVSAVFHEL CIAVP
CRLFKLWAF LGIMFQVGNMIFWFIFC IFGQPMC VLLYYHDL MNRRKGSMS"
repeat_region complement(26285..26320)
          /rpt_family="(TAAA)n"
gene      <27991..>28341
          /gene="At2g19460"
          /note="synonym: F3P11.6; predicted by genefinder"
mRNA      <27991..>28341

```

```

CDS      /gene="At2g19460"
         27991..28341
         /gene="At2g19460"
         /note="hypothetical protein"
         /codon_start=1
         /protein_id="AAD10145.1"
         /db_xref="GI:4191776"
         /translation="MEDYRSRSYGDGRSDLQQYSAHRRSDGPDSFSGNGMQDLRSYS
TSYTDYPTRIPEDQNPKKGRSSSSSSWGFVDPDLQRKKRVVSYRAYTVEGKLGKGSFRK
SFKWIKDKCNKLLN"
repeat_region complement(28413..28443)
         /rpt_family="(CA)n"
gene      complement(<30261..32722)
         /gene="At2g19470"
         /note="synonym: F3P11.7"
mRNA     complement(join(<30261..30485,30572..30660,30756..30819,
30915..30997,31093..31219,31329..31413,31493..31556,
31650..31711,31825..31895,31981..32076,32150..32298,
32379..32448,32518..32558,32630..32722))
         /gene="At2g19470"
CDS      complement(join(30261..30485,30572..30660,30756..30819,
30915..30997,31093..31219,31329..31413,31493..31556,
31650..31711,31825..31895,31981..32076,32150..32298,
32379..32448,32518..32558,32630..32705))
         /gene="At2g19470"
         /codon_start=1
         /product="putative casein kinase I"
         /protein_id="AAD10146.1"
         /db_xref="GI:4191777"
         /translation="MEPRVGNKFRLGRKIGSGSFGEIYLGTDVQTNEEVAIKLESVKT
AHPQLSYESRIYRVLQGGTGIPNMKWYGVGEDYNVLMVMDLLGPSLEDLFSYCKRQFSL
KTVLMLADQMINRLEFIHKSYSYLRDIPDNFLMGLGRRANQVYIIDYGLAKKYRDSS
THRHI PYRENKSLIGTPRYASLNTHLGIEQSRRDDIESLGYILMYFLKGS LPWQGLKA
GNKKQKYDKISEKKVSTSIETLCRGHPTEFASYFHYCRSLRFDKPDYAYLKRFLRNL
FIREGFQDFVFDWTVYKYQQSQSGNPQPRPHDGGVGTSSGLNPAVGNSEKRPDVPNQ
RTNPDFTLKQKDKNGNDSIAIAKDKLLPGSLNLRSEGSRRRVVDTSSREPFSGGSDN
ANYETALKGIDGLRINNAGDETAATPQSNNGDDVEPQSKAL"
repeat_region 34209..34328
         /rpt_family="(TAAA)n"
repeat_region 34209..34266
         /rpt_family="(TA)n"
gene      34963..37696
         /gene="At2g19480"
         /note="synonym: F3P11.8"
mRNA     join(34963..35056,35161..35204,35433..35528,35658..35750,
35973..36060,36165..36226,36309..36493,36577..36754,
36842..36885,36969..37185,37282..37299,37382..37696)
         /gene="At2g19480"
CDS      join(35011..35056,35161..35204,35433..35528,35658..35750,
35973..36060,36165..36226,36309..36493,36577..36754,
36842..36885,36969..37185,37282..37299,37382..37450)
         /gene="At2g19480"
         /codon_start=1
         /product="putative nucleosome assembly protein"
         /protein_id="AAD10147.1"
         /db_xref="GI:4191778"
         /translation="MSNDKDSMNMSDLSTALNEEDRAGLVNALKNKLQNLQAGHSDVL
ENLTPPVKRKVEFLREIQNYDEMEAKFFEERAALAEAKYQKLYQPLYTKRYEIVNGVV
EVEGAEEVKSEQGEDKSAEEKGVPDFWLIALKNNEITAEIEITERDEGALKYLKDIKW
SRVEEPKGFKLEFFDQNPYFKNTVLTKTYHMIDEDEPILEKALGTEIEWYPGKCLTQ
KILKKKPKKGSKNTPITKTEDCESFFNFFSPPQVPDDDEDLDDDMADELQGGMEHDY
DIGSTIKEKIIISHAVSWFTGEAVEADDLIEDDDDEIDEDDDEEDEDDEEDEDDE
DDDEEEEADQGGKSKKKSSAGHKKAGRSQLAEGQAGERPPECKQQ"
repeat_region complement(37036..37151)
         /rpt_family="(CAT)n"
gene      <38677..>40416
         /gene="At2g19490"
         /note="synonym: F3P11.9"
mRNA     join(<38677..38847,39014..39101,39185..39357,39458..39604,
39694..39828,40000..>40416)
         /gene="At2g19490"

```

CDS join(38677..38847,39014..39101,39185..39357,39458..39604,  
 39694..39828,40000..40416)  
 /gene="At2g19490"  
 /codon\_start=1  
 /product="putative recA protein"  
 /protein\_id="AAD10148.1"  
 /db\_xref="GI:4191779"  
 /translation="MSKKEIALQQALDQITSSFGKGSIMYLGRAVSPRNPVVFSTGSF  
 ALDVALGVGGLPKGRVVEIYGPESGKTTLALHVIAEAQKQGCTCVFVDAEHALDSSL  
 AKAIGVNTENLLLSQPDCEQALSLVDTLIRSGSVDIVVDSVAALVPKGELEGEMGD  
 AHMAMQARLMSQALRKLSLSLSTLLIFINQVRSKLSTFGGFGGPTEVTCGGNALK  
 FYASMRLNIKRIGLIKKGEETTGSQVSVKIVKNKLAPPFRTAQFELEFGKGICKITEI  
 IDLSIKHKFIAKNGTFYNLNGKNYHGKEALKRFLKQNESDQEELMKKLQDKLIADEAA  
 DKETESSEEEEDSLRVVSPDNTDDESPALVVGAAAVVVEAA"  
 gene complement(<40721..>43711)  
 /gene="At2g19500"  
 /note="synonym: F3P11.10"  
 mRNA complement(join(<40721..41012,41054..41364,41513..41770,  
 42535..42662,43153..>43711))  
 /gene="At2g19500"  
 CDS complement(join(40721..41012,41054..41364,41513..41770,  
 42535..42662,43153..43711))  
 /gene="At2g19500"  
 /codon\_start=1  
 /product="putative cytokinin oxidase"  
 /protein\_id="AAD10149.1"  
 /db\_xref="GI:4191780"  
 /translation="MANLRLMITLITVLMITKSSNGIKIDLPKSLNLTSTDPSSIISA  
 ASHDFGNITTVTPGGVICPSSTADISRLQYAANGKSTFQVAARGQGHSLNGQASVSG  
 GVIVNMTCTITDVVVSCKDKKYADVAAGTLWVDVLKKTAEKGVSPVSWTDYLHITVGGTL  
 SNGGIGGQVFRNGPLVSNVLELDVITGKGEMLTCSRQNLPELFGVLGGLGQFGIITR  
 ARIVLDHAPKRAKWFRMLYSDFTTFTKDKQERLISMANDIGVDYLEGQIFLSNGVVDTS  
 FFPPSDQSKVADLVKQHGIIVVLEVAKYDDPNLPIISKVIDTLTKTLSYLPFGFISMH  
 DVAYFDFLNRVHVEENKLRSGLWELPHPWLNLYVPKSRILDFHNGVVKDILLKQKSA  
 SGLALLYPTNRNKYILLFIHIYLPQEPKWDNRMSAMIPEIDEDVIYIIIGLLQSATPKDL  
 PEVESVNEKIIRFCKDSGIKIKQYLMHYTSKEDWIEHFGSKWDDFSKRKDLDFDPKKLL  
 SPGQDIF"  
 repeat\_region complement(47770..47832)  
 /rpt\_family="POLY\_A"  
 repeat\_region complement(48022..48135)  
 /rpt\_family="(TAAAA)n"  
 repeat\_region complement(48655..48772)  
 /rpt\_family="(TAAAA)n"  
 repeat\_region complement(48671..48760)  
 /rpt\_family="(GAAA)n"  
 gene <49640..>50086  
 /gene="At2g19510"  
 /note="synonym: F3P11.11; predicted by grail"  
 mRNA join(<49640..49740,49825..>50086)  
 /gene="At2g19510"  
 CDS join(49640..49740,49825..50086)  
 /gene="At2g19510"  
 /note="hypothetical protein"  
 /codon\_start=1  
 /protein\_id="AAD10150.1"  
 /db\_xref="GI:4191781"  
 /translation="MEPLGDRRPCVCITKNRNCPRFCEYAEYFPYELRSHYESTNEL  
 FGTPKIIKMMRHAPEEKQMLATSIIMEGNAWTNDPVSGGFGMVQKIMWKIMLHKAYL  
 HELEEKIKEEKEKIELHL"  
 repeat\_region 52348..52426  
 /rpt\_family="(GA)n"  
 gene 52414..55885  
 /gene="At2g19520"  
 /note="synonym: F3P11.12; almost identical to GB:2599091;  
 contains a Trp-Asp (WD-40) repeat signature"  
 mRNA join(52414..52699,53055..53118,53212..53298,53408..53470,  
 53644..53709,53854..53928,54066..54141,54215..54410,  
 54521..54602,54683..54756,54831..54942,55028..55102,  
 55207..55274,55359..55458,55544..55885)  
 /gene="At2g19520"  
 CDS join(52416..52699,53055..53118,53212..53298,53408..53470,

53644..53709,53854..53928,54066..54141,54215..54410,  
54521..54602,54683..54756,54831..54942,55028..55102,  
55207..55274,55359..55458,55544..55645)  
/gene="At2g19520"  
/codon\_start=1  
/product="putative WD-40 repeat protein, MSI4"  
/protein\_id="AAD10151.1"  
/db\_xref="GI:4191782"  
/translation="MESDEAAAVSPQATTTPSGGTGASGPKKRGRKPKTKEDSQTPSSQ  
QQSDVKMKESGKKTQQSPSVDEKYSQWKGLVPILYDWLANHNLVWPSLSCRWGPQLEQ  
ATYKNRQRLYLSEQTDGSPVNTLVIANCEVVKPRVAAAEHISQFNEEARSPFVKKYKT  
IIHPGEVNRIRELPQNSKIVATHTDSPDVLWDVETQPNRHAVLGAANSRPDLILTGH  
QDNAEFALAMCPTFPFVLSGGKDKSVVLWSIQDHITTIGTDSKSSGSI IKQTGEGTDK  
NESPTVGPGRGVYHGHEDTVEDVAFSPTS AQEFC SVGDDSC LILWDARTGTNPVTKVEK  
AHDADLHCVDWNPHDDNLILTGSADNTVRLFD RRKLTANGVGSPIYKFEGHKA AVL CV  
QWSPDKSSVFGSSAEDGLLNIWDYDRVSKKSDRAAKSPAGLFFQHAGHRDKVDFHWN  
ASDPWTIVSVSDDCETTGGGGTLQIWRMSDLIYRPEEEVVAELEKFKSHVMTCASKP"  
gene <56629..>57896  
/gene="At2g19530"  
/note="synonym: F3P11.13"  
mRNA join(<56629..56731,57391..>57896)  
/gene="At2g19530"  
CDS join(56629..56731,57391..57896)  
/gene="At2g19530"  
/note="unknown protein"  
/codon\_start=1  
/protein\_id="AAD10152.1"  
/db\_xref="GI:4191783"  
/translation="MESSPSGSEPPQKVVS KLQKVGWRATMIFNLGFAAYIFAIREK  
DIDADEKKKVKGSEARHKG VKKGAVNTEIEKKGAETDKAKEAETA IPEKEETKLIP  
ELDPLFEFTDATDQSMFQTVATEHVKVARKPIPEDEQKELFKWILEEKRKIEPKDRKE  
KKQIDEEKAILKQFIRAERIPKLLPDDSDVSSLRDWDKFFSK"  
repeat\_region 57729..57791  
/rpt\_family="(GAAA)n"  
gene <58214..60876  
/gene="At2g19540"  
/note="synonym: F3P11.14"  
mRNA join(<58214..58255,58373..58527,58740..58818,58898..59077,  
59186..59284,59378..59546,59635..59738,59821..59960,  
60052..60124,60209..60324,60412..60529,60623..60876)  
/gene="At2g19540"  
CDS join(58214..58255,58373..58527,58740..58818,58898..59077,  
59186..59284,59378..59546,59635..59738,59821..59960,  
60052..60124,60209..60324,60412..60529,60623..60757)  
/gene="At2g19540"  
/codon\_start=1  
/product="putative WD-40 repeat protein"  
/protein\_id="AAD10153.1"  
/db\_xref="GI:4191784"  
/translation="MGRNVKTKAKRKNKKKAEASSSEIPSIPTRVWQPGVDTLEDGEE  
LQCDPSAYNSLHGFHVGPCLSF DILGDKLGLNRTEFPHTLYMVAGTQAEKAAHNSIG  
LFKITNVSGKRRDVVPKTFGNGEDEDEDEDDSDSDDDGD EASKTPNIQVRRVAHNG  
CVNRI RAMPQNSHICVSWADSGHVQVWDMSSHLNALAESETEGKDGTSPVLNQAPLVN  
FSGHKDEGYAIDWSPATAGRLLSGDCKSMIHLWEPASGSWAVDPIPFAGHTASVEDLQ  
WSPAENVFASCSVDGSAVVDIRLGKSPALSFKAHNADVNVISWNRLASCMLASGSD  
DGTFSIRDRLRIKGGDAVVAHF EYHKHPITSIEWSAHEASTLAVTSGDNQLTIWDL SL  
EKDEEEEA EFNAQTKELVNT PQDLPPQLLFVHQGQKDLKELHWHNQIPGMIISTAGDG  
FNILMPYNIQNTLPSELPA"  
repeat\_region complement(58984..59052)  
/rpt\_family="(CAT)n"  
gene complement(<60943..>61941)  
/gene="At2g19550"  
/note="synonym: F3P11.15; contains  
esterase/lipase/thioesterase active site serine (PS50187)  
and alpha/beta hydrolase fold (PF00561)."  
mRNA complement(<60943..>61941)  
/gene="At2g19550"  
CDS complement(60943..61941)  
/gene="At2g19550"  
/codon\_start=1  
/product="putative esterase"

```

/protein_id="AAD10154.1"
/db_xref="GI:4191785"
/translation="MKPTKIVIPNRRNEKLVGLLHETGSKEVVVLCHGFRSDKTNKIL
KNVATALEKEKISSFRFDPSGNGDSEGTFFYYGNFNSEAEDDLHYVIQHLSSSNIMNRL
VPVILGHSGKGGDVLLYASKFPDYIRNVVNISGRFDLKNVDRLGDGYIEKIKEQGFID
ATEGKSCFRVTQESLMDRLNTDMHQACLNIDKQCKVLTVHGSDDTVVPGEDAKEFAKV
IPNHKLEIVEGANHGTYKHQKELVSIAVEFTKTAIVEQHNLLVVFGRWLIIVMIVVTF
YMSVFDHLIKVYPLLVIFKLENVMSRIQQLIYLIFFYLLRIKLLGLWLLLLLERRM"
gene complement(<62739..65498)
/gene="At2g19560"
/note="synonym: F3P11.16"
mRNA complement(join(<62739..62817,63153..63229,63821..63910,
64163..64252,64359..64452,64571..64707,65312..65498))
/gene="At2g19560"
CDS complement(join(62739..62817,63153..63229,63821..63910,
64163..64252,64359..64452,64571..64707,65312..65479))
/gene="At2g19560"
/note="unknown protein"
/codon_start=1
/protein_id="AAD10155.1"
/db_xref="GI:4191786"
/translation="MAYVSMGEAHRRITEYLNRFCDVASYQDSSSTLCRLLSFSSNSPP
LLSLADALNVFQDSSSLIRQSDRFSEYGEILAHVFRSLQSYRVGNLVEAYLAFDKFAN
AFVQEFRNWESAWALEALYVVCYEIRVLAEKADKDLTSNGKSPEKLKAAGSLLMKVFG
VLALGTVNLCRSVIRSIETARIFDFEEFPRDKYTKIVQALRKGDRLRLRHALQEHED
RFLRSGVYLVLKLELQVYQRLMKKM"
gene complement(<67008..67978)
/gene="At2g19570"
/note="synonym: F3P11.17"
mRNA complement(<67008..67978)
/gene="At2g19570"
CDS complement(67008..67913)
/gene="At2g19570"
/codon_start=1
/product="putative cytidine deaminase"
/protein_id="AAD10156.1"
/db_xref="GI:4191787"
/translation="MDKPSFVIQSKEAESAAKQLGVSVIQLLPSLVKPAQSYARTPIS
KFNVAVVGGLSSGRIFLGVNVEFPNLPLHHSIHAEQFLVTNLTNGERHLNFFAVSAA
PCGHCRQFLQEIRDAPEIKILITDPNNSADSDSAADSDGFLRLGSFLPHRFGPDDLGLG
KDHPLLLESHDNHLKISDLDSICNGNTDSSADLKQTALAAANRSYAPYSLCPSGVSLV
DCDGKVYRGWYMESAAYNPSMGPVQAALVDYVANGGGGGYERIVGAVLVEKEDAVVRQ
EHTARLLLLETISPKCEFKVFHCYEA"
gene complement(<68803..>71431)
/gene="At2g19580"
/note="synonym: F3P11.18; contains a transmembrane 4
family signature; rare (GC) splice donor consensus found
instead of (GT) at intron 2."
mRNA complement(join(<68803..69114,70846..71152,71238..>71431))
/gene="At2g19580"
CDS complement(join(68803..69114,70846..71152,71238..71431))
/gene="At2g19580"
/codon_start=1
/product="putative senescence-associated protein 5"
/protein_id="AAD10165.1"
/db_xref="GI:4191796"
/translation="MALANNLTAILNLLALLCSIPITASGIWLASKPDNECVNLLRWP
VVVLGVILVVSATGFIGAYKYKETLLAVYLCCMAILIGLLLVLLIFAFVVTRPDGSY
RVPGRGYKEYRLEGFSNWLKENVDSKNWGRRLACLADTNVC PKLNQEFITADQFFSS
SKITPLQSGCCKPPTACGYNFVNPTLWLNPTNMAADADCYLWSNDQSQLCYNCSCKA
GLLGNLRKEWRKANLILIIITVVVLIWVYVIACSAFRNAQTEDLFRKYKQGWV"
repeat_region 71243..71300
/rpt_family="(CGA)n"
gene complement(72513..73785)
/gene="At2g19590"
/note="synonym: F3P11.19"
mRNA complement(join(72513..73234,73335..73555,73639..73785))
/gene="At2g19590"
CDS complement(join(72649..73234,73335..73555,73639..73764))
/gene="At2g19590"
/codon_start=1

```

```

/product="1-aminocyclopropane-1-carboxylate oxidase"
/protein_id="AAD10157.1"
/db_xref="GI:4191788"
/translation="MVLIKEREMEIPVIDFAELDGEKRSKTMSSLLDHACDKWGFFMVD
NHGIDKELMEKVKKMINSHYEEHLKEKFYQSEMVKALSEGKTSADWESSFFISHKPT
SNICQIPNISEELSKTMDHEYVCQLHKFAERLSKLMCENLGLDQEDIMNAFSGPKGPAP
GTKVAKYPECPRPELMRGLREHTDAGGIILLQDDQVPGLEFFKDGKWVPIPPSKNNT
IFVNTGDQLEILSNGRYKSVVHRVMTVKHGSRLSIATFYNPAGDAIISPAPKLLYPSG
YRFQDYLLKLYSTTKFGDKGPRLETMKKMGNADSA"
repeat_region complement(74164..74272)
/rpt_family="POLY_A"
repeat_region complement(75620..75652)
/rpt_family="(GAA)n"
gene <75730..>79892
/gene="At2g19600"
/note="synonym: F3P11.20"
mRNA join(<75730..75952,76149..76192,76277..76341,76441..76504,
76623..76684,76792..76867,77020..77094,77224..77295,
77387..77470,77579..77647,77830..77886,78105..78193,
78325..78421,78501..78554,78656..78763,78945..79133,
79208..79285,79431..79475,79635..79724,79809..>79892)
/gene="At2g19600"
CDS join(75730..75952,76149..76192,76277..76341,76441..76504,
76623..76684,76792..76867,77020..77094,77224..77295,
77387..77470,77579..77647,77830..77886,78105..78193,
78325..78421,78501..78554,78656..78763,78945..79133,
79208..79285,79431..79475,79635..79724,79809..79892)
/gene="At2g19600"
/codon_start=1
/product="putative potassium/proton antiporter"
/protein_id="AAD10158.1"
/db_xref="GI:4191789"
/translation="MRLTLTLLICTFFFFFSFAYSAESDNETDSVVTREINGTVVESN
ATSAKPREDSFADMIDRALEKEFPDNDQNEVPDPGSFNNSVADQQAVLETVARVKPKK
NETKTKEEKSFFNLDNENGVEDTPRLIDRKDNVFI MSNPKSKYPVLQDLRLISDLVV
VIVSATCGGIAFACAGQPVITGYLLAGSIIGPGGLSFVSEMVQVETVAQFGVIFLLFA
LGLEFSAAKLRVVRVAIPAIGLLQIFLFMCLSGITASLCGGKLTEGIFVGAFLSMSST
AVVLKFLMERNISALHGQITDCAVGLLFALLPVLGGTSGVLQGVLSMAKSLAILIAF
LGALFVLSRTWVPWFLKLMTSLSSQTNELYQLAAVAFCLLVAWCSDKLGLSLELGSA
AGVMISTTDLAQHTLEQARSQVEPIRNFFAALFLASIGMLIHMHLWNHVDILLAAVL
LVIVIKTVVVAIVVKVFGYNNKTAVLVGMSLAQIGFAFVLLSRASNHLHIESKLYLL
LLGTTALSLVTTPLLFKLIPAVVHLGVLLRWFSPDSSTEIGFKGELYHSESAKRISLM
IQGSLHDS"
repeat_region complement(75736..75791)
/rpt_family="(GAA)n"
repeat_region 80290..82654
/note="putative retrotransposon; Structure: LTR,
zinc-finger gene, LTR."
repeat_region complement(80402..80516)
/rpt_family="(TAAAA)n"
gene <80865..>82133
/gene="At2g19610"
/note="synonym: F3P11.21"
mRNA join(<80865..81272,81348..>82133)
/gene="At2g19610"
CDS join(80865..81272,81348..82133)
/gene="At2g19610"
/codon_start=1
/product="putative C3HC4-type RING zinc finger protein"
/protein_id="AAD10159.1"
/db_xref="GI:4191790"
/translation="MDNHLELAIKDAITADDLKRVDQETQHPLLAQELDLDSLENPPR
ATHTYRLYSKGLVSEELIKDDTMLVVGLGLSLCDSHDYTKQEINKALRNQKLAHPE
AAELAIIHGLKWALELGIERIQFFCDDSNILAYVTRKAAPNESIVAKLLEHVSLTQT
RFTSCQALATVSRDDIVSVIKLAKDAIASQTRWCEGDTEYESCPVCYAYVSPNDKFEV
QGC FHRICVTMRKPFSSSEQILRGNTAICPYPDCENDLVPEDCRAFADADAITLMIQR
KKEKAIPVKDRVYCPNPSCSFLMSDLDLIRHISKPNRHSEEARCMECGLSFCKKCHV
PWHYKKTCDFFKSESYLKSDAAILES FVKQTQGWKKCSQCQSIVQHGGGCQQMTCRYW
TFVIY"
gene complement(<82541..>85061)
/gene="At2g19620"

```



```

mRNA      /note="synonym: F3P11.22"
          complement(join(<82541..82762,82873..83064,83161..83217,
          83298..83375,83461..83547,83654..83714,83818..83906,
          84014..84111,84211..84310,85002..>85061))
          /gene="At2g19620"
CDS        complement(join(82541..82762,82873..83064,83161..83217,
          83298..83375,83461..83547,83654..83714,83818..83906,
          84014..84111,84211..84310,85002..85061))
          /gene="At2g19620"
          /codon_start=1
          /product="putative SF21 protein  (Helianthus annuus)"
          /protein_id="AAD10160.1"
          /db_xref="GI:4191791"
          /translation="MVGLNNAVSLDIEEICNGGKEHHVKTCHGSVSVVVYGDQEKPAL
          ITYPDVALNYMSCFQGLFLCPEAVSLLLHNFCIYHISPPGHEFGAAPVCSNDPSPSVE
          DLADQILEVLNFFSLEAVMCMGITAGAYILSLFAIKHKERVGLILISPLCKAPSWSE
          WFYYKVVSNNLLYYYGMSGLLKDIFLQRYFSKEARGSSSEVPERDVVHECRRLGERHGS
          SLMRFLEAVNRRHDLTDGLKSLKCRTLIFVGDQSPFHSETLHMVTALDRKYSALVEVQ
          ACGSMVTEEQPHAMLIPMEFFFMGFGLYRPGRVSDSPRSPLSPSCISPELLSPESLGL
          KLPKPIKTRVPTKC"
repeat_region 85525..85620
          /rpt_family="(TAAA)n"
repeat_region complement(85675..85780)
          /rpt_family="(TA)n"
gene        complement(<86627..>87520)
          /gene="At2g19630"
          /note="synonym: F3P11.23; predicted by genscan and grail"
mRNA        complement(<86627..>87520)
          /gene="At2g19630"
CDS        complement(86627..87520)
          /gene="At2g19630"
          /note="hypothetical protein"
          /codon_start=1
          /protein_id="AAD10161.1"
          /db_xref="GI:4191792"
          /translation="MKTSLRLEDGTKNSLQIPIDLIIEIFLRLSVNSIARCRCVSKQW
          ASTLSRPYFTELFLTRSLARPKLLFAYRKGS DYFLSSPQLQNPDDDHKKSSPVVNY
          HMHHILTLGSGNMSWRTIQCCIPHRSFDDGGICIDGLIYYYAGVNNNDIVIVCFDVRSE
          EFRFIIGALHHGSLINYNKLSLYLPSTVYSRFGVIRSILKWLEDAKRQEWSEHTY
          ILPAMHDIMKTDLRCVGMTQTKEFVFWFSIKGMRFYVFFYNIERNITIKKVMQGMFAF
          KESNVYTFLDHVEDVKLMQFF"
gene        <87811..>89007
          /gene="At2g19640"
          /note="synonym: F3P11.24"
mRNA        <87811..>89007
          /gene="At2g19640"
CDS        87811..89007
          /gene="At2g19640"
          /codon_start=1
          /product="putative SET-domain transcriptional regulator"
          /protein_id="AAD10162.1"
          /db_xref="GI:4191793"
          /translation="MINDGGAKPETLLRVAEIGGRGRSLVAAQSLRAGQVILRESPLL
          LYSAFPFLSSSVSPYCDHCFRLASSAHQKCQSCSLVSFCSPNCFASHTPWLCESLRR
          LHQSSSSAFSDQPSDRQVQARFLLSAYNLAAASPSDFQILLSLQGSGSSNGDPSCSAG
          DSAAAGFLHSLSSVCPSPVPVISPDLTAALLSKDKVNAFGLMEPCSVSNEKRSVRAY
          GIYPKTSFFNHDCLPNACRFDYVDSASDGNTDIIIRMIHDVPEGREVCLSYFPVNMNY
          SSRQKRLLLEDYGFKDCDCRCKVEFSWSEGEEDENEIMEEMEDQDEQEEMEDSVGENEE
          EVCGNGVDDSNFPHAYFFVRYMCEKENCFGT LAPLPKTHDASRVLECNVCGSVKED
          EVGVNQ"
repeat_region 88720..88816
          /rpt_family="(GAA)n"
gene        complement(<90065..>92113)
          /gene="At2g19650"
          /note="synonym: F3P11.25; similar to GB:AAA75253; some
          members of this protein family have a weak CCHC zinc
          fingers that is mostly from retroviral gag proteins
          (nucleocapsid)"
mRNA        complement(<90065..>92113)
          /gene="At2g19650"
CDS        complement(90065..92113)

```

/gene="At2g19650"  
 /codon\_start=1  
 /product="putative Ta11-like non-LTR retroelement protein"  
 /protein\_id="AAD10163.1"  
 /db\_xref="GI:4191794"  
 /translation="MENFLALFNEGMDDEEASRLEKLLPFHEHPLIPFTRFEFTKCKGC  
 YSTDEDNTSFISKHPGCSYIYGGHRCNEPGCDVVFHEECAKPLPEIDHPYHPVHPLKL  
 ILLHTRVSNCSFCGEFFSVGYRCNKCDFQLDLICARRPALLVLAENSYLHEHPLQLSH  
 EAHETLRDCKGCGHQHNSWKHFYTCHQCEFFISLDCIDHNSPEAYHTSHPQHPLKSLR  
 YEVPGYADKKCLLCREEFDGIHRQLHHCVCNFSICRSCMKNPVVLVKSPTHEHQL  
 HLVPRRIDFTCNACGTGLDRSPYFCLQC�FMHRECIDLPRVININRHDHRI SYTPRL  
 GHGDGKCRVCRKNVDGFYGA YSCSKCFDYFVHSRCATRRD VMDIELEGTPEEEETAP  
 FKVIDDNTIDHYGHKHNLRIRGWNL PESV VCTACVFQICSEPFYACARKRCDFIHE  
 KCAKPLPKKRHVCHNQPLELSPNYGTYLLHRC SICRQLFTGFSYESKELRIDVRCATV  
 TEPFEHKSHQHPHYSTDH SKRCSECDGDGKLS CDDCEFGLCFKCAILPQKVMRHR YD  
 DHPLFLSYGEMGVAGKYWCEACETT VNP EKWFYTCNECGITLHISCVIDTYAYSMPGA  
 LTWIGTFVSNHSICRPLCCVCNSRCKLPFVRTLPEDGADLYFCSLECEKKIYQERFEF  
 "

repeat\_region 90954..>92721

/note="putative retrotransposon; Structure: LTR,  
 zinc-finger gene, LTR."

# ORIGIN

1	aaagggccac	aacttgtagt	acacaaaatg	taatagtaga	cttcatttgc	ggacatgcat
61	caccagtgtt	ccaaaactgt	caagttgagg	cctgaaacgt	gacatagcaa	tgttataatt
121	gttctgtcaa	tggaaggaca	gagcaatgtt	ataactgctc	aacaacgtga	taagtctggc
181	gaaagtgttg	gattttcaat	tcaaaaatgt	gaaaatatac	taaatatgta	ctcatatact
241	aaatatacta	aaaagtattt	ttgtgaattt	atgtgagttt	atgtatgtta	acgtgtatgg
301	ccgtttgata	tggaagtgg	atgtttat	ttggccaagg	gagatagtag	tgttgcta
361	attgtttttt	tgaacgactt	gatccctcat	tgaaggttta	gttctatatt	aaacttgtct
421	gtaatttgaa	gaagtcacg	agcaaagcct	tgggtggctt	tgtttgtagt	tatctttggt
481	tggtttggtt	aaatgatttt	taaaggta	atgttgga	gatcaatgga	tgtgaaagga
541	tgatcaagaa	actatgttca	tgaaaat	gaaat	ttcaact	atgcaaaaga
601	tgagcagaat	tacgatcctt	tcaaagcctt	ttttttcac	caccatattt	tctctcagct
661	attgtaaa	gattggttta	gttacatttc	acgtaaactt	tagtgataat	caaaaatattt
721	tatctta	acattagacc	cttgactttg	tattgtatct	cacatttttt	cttctagtta
781	tttcttaggt	tctttaaaaa	ataattactg	tatttaacta	aattagatat	gccatatttt
841	atttcttggt	ccaagtacca	aagagcctag	gcacgacaat	gtgtgtat	actctttaat
901	ggttctataa	caaaat	gtgaaaatgg	tggtgtttgt	tttaggctat	ttatcaatat
961	cttcgagggt	catgattcta	tttgttat	cat	tttctgata	ttattctgcc
1021	cccattttgc	accac	tcatgaggtt	tgaat	atccattaga	aaatattatg
1081	aaaacatctt	tgacacaaaa	aattatgaaa	ccaaatgtaa	tattcaaca	actaaaat
1141	tcatttttat	tggtttacaa	acataatcca	aaccacgtac	gtaactctct	tttttggtt
1201	atcacataaa	tatcacaatt	tacaaac	acatttacat	gcatataatt	aaaaatattg
1261	caccttcaaa	taaatgtttt	tacaaaaatc	cccaatgaga	aactagccaa	ttaaacaacg
1321	acattagaaa	accggactct	tgatttccaa	tcgcgaatcg	cagaccggag	cgtatgggtc
1381	ggtgtgagct	tgaagtggtc	tagctttt	cgagtcaccg	gtgaaatgtt	atgtttctca
1441	agccacgcta	agatcgctt	cctctcgtaa	gtgaatccat	cggcagcaat	ctctggctcc
1501	tccattatct	cctacatcca	acaatccaaa	taaataaata	aaatctaaac	cggttcaaat
1561	ttcttggtac	tacatttg	taccgaacca	ggtaac	tactcaccg	taagattgga
1621	cagaagta	gactcgggtc	acgtaaat	cttccttctt	tctttacttt	cgaattcgcc
1681	gtctccacaa	gccgtttcaa	aaccgggtata	acctctgatt	taagatccgg	tctatctcgg
1741	catcgaaatt	cagcacattt	taaaccgatc	cgagccaa	cctcgggttc	tgccaaaggc
1801	caatcggtaa	ccgatttatc	tagcatttcg	gttaacgtcc	ctttctttac	cgcgttttcg
1861	accgctggta	caatcccgct	cgggttacga	gccgtcaaca	actgaagaat	gattatgccg
1921	aacgcatata	gatccgactt	tggtctaatt	gttcgggttc	tgtcggtattc	gggatcaata
1981	taatgcaatg	taccggcaag	aaccgagttt	cggtacatcg	tgacattatc	gggtgcaaca
2041	tccgtaacca	gcttggttag	gccaacgtca	gcgattttgc	taacgtagtt	cgggtttaac
2101	aagatatttc	ccggttttaa	gtcacggtga	acaattggtt	ccggtttaga	gctgtgtaag
2161	aaggccaaac	cgcaagctac	ctcga	accctaaacc	ggataaacca	aggcaaaggc
2221	ggtttat	ttcggtgaaa	tatatattcc	tcgaggcttc	cattttccaa	gtactcgtaa
2281	accaaacaac	cattctccgg	acaagctccg	aggaggagaa	ccacgtgtgg	gtgtcggagt
2341	tggcttagaa	cctcaaccta	tagagaaata	ttaccaaacc	gaaaataaat	caaccggatc
2401	aaatcaaaag	agttaaccga	aaccagacat	aat	atacctcttt	caagaactct
2461	tgtttcttct	ccggtgtatc	tagccgggaca	actttaacag	ccgcgggagt	actatcaagg
2521	ctacattggt	aaactttccc	gtatcctcct	tctccaatca	ctttctctgg	cgagaatcct
2581	tctgtggctg	ttacaatttc	ctcaattgtg	tattttctgt	accggtgatc	agtccctaaa
2641	agctgatcga	tcactttctt	cttctccaag	taagtcctta	aagcattcac	ctccgcgatc
2701	tgtcgtgac	aaaactctct	tgcgagcaac	gctttcgcgg	tttcgacctc	tttcacggcc
2761	ttcatgtacc	gctctttctc	cagcgccg	gtgttctctc	gcaactcttc	tttctccaca
2821	gcattgttca	ctctttttaga	ttcattcaaa	tattcagtcg	aaagcatttt	aacctaacag
2881	aaaccccaa	ataaacactt	tacttggtgaga	aaaagcaagc	aaaatgattc	accagttagt

```

2941 ttggaaaact attttataga aacaataaaa tcagatataa aataaacact tctttcacag
3001 ataagttttg atggccatca caaatatata aatttctctt cccaaaatca tctctaata
3061 ttttcttgca gatttttaggg ttcaagatca gaaatctata ccttgttttg tgtggagaag
3121 agctcttcac aagcttgttt atattttaaag actgtacttt gtactcattt cttcaaacgt
3181 tcgacctctt cttcaatctc aacctgtcaa acaacaccaa atctttcaaa aacctcggtt
3241 acgtcataat agctagtctc atcattgacg caaaagatga tggtaaagta accctcttgg
3301 attttcttga tgtttccggg ggtgaacgat cagagtctcg gtgttcgcta acgtcttct
3361 caacatttga ttggggcttg gtacatgttt tatcaaaatc tgagtaattc agctgaggaa
3421 tatcggtacc tctccgtcgc ctaaaaatct ctggagtgtg tgcactagaa gccttactac
3481 tttgaggcgt tttgggttta ttacaagtaa ggcttaatgc ctcaaactct agctctcttg
3541 ccgatgtcga cctccttggt cctgcctctt caacatccaa aagaaagaga aaaacttcag
3601 gcttcaaga catcatagtt tatattatat gcattatcat ttatatattt tataatatga
3661 atctatactt gatagatgca tgctcgttgc aaatttcttc gtataacaaa tttatatcca
3721 aagttataat gttatgcaac tagtttactt tttgttaact aagattgttg ttatatattt
3781 ttcaaaatta ggttcaattc aaaaaaatga gtcaattgac cattaaaatt atttttattt
3841 ttgtttgttt aagaatgtca gaattgagaa tgtactatat attaaaacaa cctaaaaccg
3901 ctattcattt cctatgaatt ctgattgata tttgactagc tatatgtata ttttaccagt
3961 tgactgtcga ggatcgggta acgttgagaa ccgtagagtg tggaaactag ccgccagtc
4021 tctcaagaag tcatgggcag tcgcggctgc gtgtggactt gtgcacggct ctggttacca
4081 cgattaaaga gaagttttta gaaatgagaa aatattgtgt tatcaaataa aatgttttgc
4141 ctataaatta ccatgtttta gaaattacat caactatgga tgtagtatta aacaataatt
4201 ttcttttaaa aaaaaacgaa acacaaaatc ctaattttta cctcggttga ttaattggatc
4261 catagatttt gtagtgattc tgtctttaca cacaatgtat acttcacatg tttctgggtc
4321 gtatcttaaa acggtcagtg gtactcctgt accttttggt ctctagacc gaaaatatcg
4381 ctggttagtgt atataaccga gttaattaag aagaacgtaa accaaaaccg aactatgtaa
4441 atttaccatg tgaatatggt tgaagtaaat gatcccataa ccaaactggt aactcctgat
4501 ttggatatca atctgacaag atattttcag atattttcag gagtttaaag tttgtctctt
4561 accttactac gaaggttaca acaaaaacaa ccttgggtact cctagtactc ttgcacattt
4621 gatttataaa ccgaaataaa gtttcatatt cttttttcac gtctcttaca tacatttcca
4681 tcaaaaaggg acaaaagact ctccacttct tccaccggca atctatctcc ggctgtaaca acccacattc
4741 ccacactctc taatatcaaa atattcttta attacaatga taaattaagt atttaaaaag
4801 tcgtaaacad atgtcaaaca aaaattttga taagtcgaat tagctaattt actttttaac
4861 atgtcaaaca cgaacaatt tcacctcatt aaaatcaaac tacgaaaaag taattactta atcatttgat
4921 cggaacaatt acataatttg taagatttgt gaaacttact tttcacattt tcttcttcca
4981 tatcatcgca atttgggtgt aaaaaaaaac ttcagtaatta tttattctct
5041 tgttctatat ggtttttact tcattttatc tttatatcat ttacattagg tttttaataa
5101 tctgtctttt aaaaatgtga atgattgaga gagacttact cggagttaga atagaagtaa
5161 caataacaag gacgtggatc atcacaaatt tatcagcctt cggtaaaaga ttatcgacag
5221 tggtagggat ggcacgacga ctctcctgcac cgctaattt accaccaccg ccactcattt
5281 cccaccgtac cacaacacgc tgacctctt cctacttac gtataccaca agtaaccaaa
5341 cggccaccgc catcaccacc atctaactca ccttgtgagg ctgatacact tttgtagtgt
5401 cttcgtcag ttgttgttgt gactaggaaa tccagagatg gatattactc tgatcacgac
5461 gaataacttt tttgggtggg gaaatgctat tccagagatg gatattactc tgatcacgac
5521 tgtttctgaa tttgggtggg gaaatgctat tccagagatg gatattactc tgatcacgac
5581 tatattgtat tgatctcttt atatatatat atataagagt cgtttaggac caagtgttta
5641 aaaatgtaaa aaaaaaaaaa tttgattagt agccgaacaa tgattgggtc tctttcggtc
5701 tatttgttgc tttggtttgt ttaatcaata attaaagtcc taagatatgt tttaatataa
5761 acttgaacaa acaaaatata tgagtttgat tgtagaaatg ctttaattaa cactctaagt
5821 ctaaaagtata aacaaaataa caacatcaaa aactaagtga ttgtagtgtg ggcgtaataa
5881 tattttgata gtccgttcta aaattagttt ataaatgcta tttttttatg gaccactatt
5941 aattattaaa tctcaaaaga attatttcta aacaatgttg tttacttaag tgatatagt
6001 cagtacaata taaaattttt cgtattagtt tataaataca tagcttaaaa aatataagag
6061 aatacgaagt agactcatag gattggatta tataaaatca gtctattgac ttttaggaaca
6121 aacaaaagat ataatttgat ctgttttgaa ttttgttgg aatcaccgat atggatcttg
6181 cagtccgtga tatataatta atagtaaaat atgtttggcc aacattagga agaaatcgat
6241 gggatggaaa tcgccttcaa aggaatatat attccattta atacttttca aattcgaatt
6301 tttttggacg gccaatttgt taattataat tagtgtgcag ctttattatc tgaagtttct
6361 tagttctatc tacatgatat ttgaacgggt tagatagaga tttttatgtc tgcagtgtt
6421 ataatatgta tcatatacta tgaccagtgt taccaatgtg ggcaatgtga tcaatataat
6481 gtaacctact aacttataac ctatgtattt gttgcaaaat aattatgtat gaagtaattt
6541 tgaatttatt tgaatttttc cctgactttg tccgtgtcaa caaacaattc gaaatgcctc
6601 gactttttta gaacaggttc tagcaaaata acttagctca gtaagctttc agaataaata
6661 aagtagtcat tttctatcag aagatgacat aatttagttt tttctctgtg gcaaacaatg
6721 acataacaaa aaacaaatga cattacttag ttacgggtact agattaatct tgaagtggta
6781 tgtggtcgca acaatctgaa tcttttagta aaaggcatga gattgtgtgt gcttgtgag
6841 ataagtcaca ttttgttccg gttaaaatat gaaattatc agtatttttt gtcgcagacc
6901 aacctatata ttaaaattct ttataatttt tttttaatac attgtcttga atttccacga
6961 tttcttgcac ataatagttg tttaggattt gatcatctga tacggtaacg tacagatcca
7021 aatttccgat caaacgggtt tgatttggct gagtaatgat gtttgcaatt gttttcta
7081 atgtaaccaa aagttgacca aaacagctca aatgattttt catataattt ttgctcagca
7141 caaatgagga tatacccgaa taattaagaa ttatatataa aaaaaaaca ctaataatgc

```

7201	actgatcggt	tacgagaacg	acataattaa	actaattaaa	gatcttaacg	attgtcgact
7261	ctccaatggt	cttatacaca	cacaaatata	aatatgatta	gtccaatcta	tatagtcgaa
7321	cagtaattta	cagcagactt	agatcgatct	ttcgtataga	agaaagagac	gttaccaatg
7381	acattactga	aataacttct	tttttttttt	taactacata	gaattattaa	cacattctaa
7441	tacagcatat	attaatacat	atatactctc	actacgtgta	aatgatagaa	gacgatcaat
7501	attgggctaga	ggtccatctt	tgggttaattg	ttccatataat	gtagatctaa	gctatatata
7561	cataataactt	tccggaatttt	tttgtattttc	ttaaaacatt	aaatatgcaa	atgtaaacgg
7621	aaacaaatca	aacaaaactt	ttaatttgat	gtcaagaaaa	cgatatcata	tttttttatt
7681	ttagcccaag	ctgtttttgg	gtgttttatta	gtatacaatt	tatatattat	tggagcgaag
7741	cctacatata	gaaagttag	agcatcatgt	agacgtcatg	tgaattctag	gtccaaaatt
7801	atgtacacac	tacataaaat	atcataacat	caaacgaaaa	acacatatac	cacacacgaa
7861	aaaagggaaa	aaagaaaaga	aataaccacc	atcacttagt	ttctgattct	ggacaagttag
7921	ttcatgaatc	atctaataata	ttagtcatta	ttgatttctt	aaagatttat	tgatcaattc
7981	atacatatat	ttctattttct	agccaaatat	ataatagagc	tagattttta	ttttttttgt
8041	tcataaaaaag	aaaacagctt	taattgataa	cagttgttca	ttgttatttt	tgctacttag
8101	acgatggatc	atgcatgaaa	aggtaccaag	taattcaaac	agtactgtta	ttggttagatt
8161	ttagatatata	gcataaatgt	gggggacaaa	aaaaatatag	atacataaat	ataatggaaa
8221	tggtaaaaga	acaaaaatag	ataaaacata	gataaaggga	gaaagaaaa	tactttttta
8281	cttaagaata	tatttctgtc	cacaacaaga	tttgctagga	cgaaaaatca	gtaactattt
8341	ctcaatttctt	ctatgtagct	ctttaaaata	ctaaagggtca	tccgatgacc	ggcatcaaac
8401	ccttatccgc	aacttgatac	gcacgctttg	ttcctttgcg	atcatgcgtg	aaatcttctc
8461	attctcatga	tgataaacct	tccttatgtc	ttgctcctta	atttttttct	ttgttaaagt
8521	gccttgctcc	tccttgtaac	tcaaagctca	tcatttgtag	aaaaattaga	tcattttgtg
8581	gcctttatta	tggtatatata	attgatgtga	ttagtggttg	taatttgata	gaaaaaacgt
8641	acgactctga	gctttacaac	ggctatttgg	caattttaca	cggatattggc	gacttacgct
8701	atttactgt	aaagattgaa	aatgatattt	gtgtgtcggt	aggtaaaaa	tacttgaaaa
8761	caaaagggtg	atagacaaga	tatggctaac	tctatctaac	catacacaga	gacgtttgct
8821	gccatgcatg	catgatgcat	atatatatata	atatgtggat	ataaaatata	tttaatttat
8881	atacacatag	tccatagata	gatagataaa	tagacgtaga	gagggagaga	gcaggagaca
8941	agaagagagt	aagcacgcag	agctttcgag	catgagcatg	ctagaaaaag	agccatgtgt
9001	gctcactctc	ttctgtcgcc	ttcgttattg	ctattccctt	acctccttta	gttgacttgc
9061	ttcatcttct	tgcatgcata	tgattttgat	gttatattag	cctcgggtccc	taaaattact
9121	atatgatatg	atgatatgat	atgaaacagt	ttgtctgatt	ttagggtcta	tatataagga
9181	ttgtcagggt	ctatatatga	tgatgcatgc	atagatttca	tcactcatat	ccactgacca
9241	ctagctacgt	ttaattttata	gaaagattcc	gccggattaa	actcttattt	cctgtcaatt
9301	tgagcaacaa	gattctctct	tttcaattat	atatacataa	cacatgtcta	taataatcgt
9361	acaaaccaat	aaaattggac	tcttaacaaa	attttgagag	tgagagatcc	cacttaaaac
9421	gaaataaact	atttgtttgg	tcattgcatg	cgtaaacgac	cacttgaaca	attaagtata
9481	cataacatag	gtctagatct	atgtgaaaaga	ttcccaaact	taagatctct	agctagaaca
9541	ttcattgatc	ctccaatcaa	aaaccaaagg	aataaaaaagt	aatctttgta	agggcgaatg
9601	taaaaccgaa	acgtaagcta	gaggaattga	tattaaaatg	gaagtaaaag	cgaattattac
9661	taaagggttt	ggagttaatt	aaataaagac	aaaagatgca	aggacaacga	attatgaaca
9721	tcttcgaaat	acgtttttta	tattgttttg	gcatttgatt	tgatttgatt	ctccttccac
9781	attcccttgt	tcatatgtgt	ttgtgtctca	cgcaagcact	acataagcgg	ttcccctaaa
9841	tacgttttcta	gaaacgtcaa	cgtttattca	atcataccaa	aatttacatg	tgtattcatc
9901	aatcagccga	ccataaaacta	catgggtccat	tagtcattac	ctatactgca	atttttctag
9961	cttttaacag	ccattatgta	cttcacagtt	tggaatactg	gtatgacgaa	gaaaagtcca
10021	ggtaaaaaaac	gtattagaat	aaatgtttat	cgtccaaaga	tcgcaccagg	taagaataag
10081	aaacctaatac	caacaaagga	acaattaatg	gacctgaaat	tttctgatga	agacgtactt
10141	accagcttgg	gttttgatga	tggcgggttac	ataataccac	taccgaaaaa	atctctcatg
10201	ttgattaatc	taccagattc	ttttcgagat	gagttgcaaa	cgatgcatgt	cagtttttat
10261	ctccgagaag	ttggcaaaact	tatgtttggg	tgaggacaaa	tatgctctag	tgaatagtac
10321	taagaagttg	atgcttacag	atgggtctagg	acgggttcat	ttgttttgga	tgaattcaat
10381	tttttttccg	aatcagttat	tgccaatgag	atttatgtag	ccctagatgt	gagatttatg
10441	aaagctggaa	gtgccaaaaa	gaaagtgga	gatttttaatt	gtaacaaaaa	cgcacagagt
10501	cttctttatg	attttttctt	ggtagatata	gatgaggaga	tatatgggtc	aaaaattgaa
10561	gagacggaag	acaaaaagaa	gtggaacaaa	aaattttgaa	ttgatttgct	ttcatcatat
10621	tgtgctatta	acattaaaca	cgaatatagag	catctgaatc	ttgaattcta	tggaatctct
10681	aaaacgaata	gacagatctc	cgcggaggaa	gtgaaattct	ctactcatac	tgcatcctaa
10741	attaattgag	tttttcaatt	acattggagg	tttatccttt	gatactttcg	gtttgggtgat
10801	ttgaatgaaa	tggttggtttt	gacagaaaaa	tgaataaacac	ttgcattaaa	atcatcaaaa
10861	attgaataac	actagaatct	aaggagggtg	attcgatctt	atatttttaag	ttatttgagt
10921	ttttatataa	ttttgatttt	ttaaagaatt	ttgatgaatt	taaggagggt	gttatgattt
10981	atggtaaaat	cctctcaaat	tccacctaaa	accatgagat	ttaaaattat	atatttttaa
11041	ctaagaaaaat	ccacttaagt	ctcttaaagt	cactaaaatc	caatctttaa	aatgttttca
11101	ataacagaag	attttttaagg	agatttttaaa	atcagtaatt	caatagcatt	agatttcatg
11161	atacttttaa	aaattcataa	ttgaataaca	taggatattg	tagtttgaat	aagatatcct
11221	aggtgttaat	ttactatttt	taaaaaggta	agattttttt	tcattgtttt	aaaaaatgtg
11281	agagtttttt	tacctatttt	tctattaaaa	tgaatgtat	aaaattaaaa	ttttgttttt
11341	aaataaaaatt	actaaaattaa	attattacaaa	atatgtatat	attactattt	agaattcaaa
11401	tattatcgat	acaaaattaa	atttatttct	tcaattcgtc	tatttaagca	agctagtaat

11461	ttgatcgaca	aaaaaaaaac	aagctagtaa	aatataatta	acagaattac	taaatcggcc
11521	gggttaggga	ttgctttctc	tcagtcgtca	attgcactaa	accactgggtg	gtaatatcgc
11581	tggcggtttcg	caagtatcac	cgggaattgc	aaagaagcag	aggaccatgt	acggagacgc
11641	tacaaactgg	aatgaagatg	agtatagaga	atcaattttg	aaggagcgag	agatagagac
11701	acgcaccgct	ttcagaaccg	cctgggctcc	tccggcgaga	atctctaate	cagacgcatt
11761	tgtttaggcc	tccagcgatg	gaacttttagc	tttccattca	ctgaactcgc	ttgtgtctca
11821	atcggcgagt	tttggctact	cgaagggtgc	agatgttatg	gtggctgaac	ctgagagagt
11881	ggttagggca	cacgaaggctc	ctgcttatga	tgtaaagttc	tatgggtgaag	acgaagatgc
11941	tttgctactt	aggtagagac	ttaaactctct	ctttgttgat	tgatgagtct	ttgatgaatt
12001	gtaattgaga	gtttgaaatt	ttgatctggg	tatgtgtagt	tggtgtgatg	atggtagagt
12061	taggggatgg	aaatgggagag	aatttgctga	atcagatgtg	tctcttcatt	tgaagggtgt
12121	gtgtgtgccca	tttttgtggg	ttttgatatg	aatgatgctt	tttggagcct	agtggttgtt
12181	tttcttatga	tctacagaga	atcatctgaa	gccattgctt	gaactgatta	atccacaaca
12241	caagttagta	tgagttttgt	tcatttttta	tttacttga	taactctgtt	actgaaaatg
12301	ctataattgt	gggttcaaat	gtcagtgttc	gtaatagtgt	tggtataaat	cactattctt
12361	tttcttatag	aggtccttgg	gggtgcgctt	caccgatgcc	tgagatcaat	gccatgtctg
12421	ttgatcctca	ggttggaatt	gtattttaa	ttcattgctt	ggtagacag	tctcacagga
12481	taactgtttc	accctctctc	acttctttcg	ttttgatagt	caggaagtgt	atctacagca
12541	gctgggtgatt	cttgccgata	ttgttgggac	gtggatgtga	tttgaccatg	atagatgtat
12601	gtccttatgtg	aatccagata	tcttaacttt	ctaaaatcat	ttatgtgtat	ccaacaattt
12661	caggagagtg	gtaagattaa	aatgaccttt	aaaggtcatt	cagactattt	gcatactgta
12721	gtttctcgta	gttctgcaag	tcaggttaatg	atgttaccat	aaatattaga	aatgcacctt
12781	tccaaatctg	agttccacat	ttggattttt	gtctgattgt	agtctctttc	tccagatatt
12841	gacgggttca	gaggatggga	ctgcgagaat	ctggggtaac	atgctcccga	ccatatttct
12901	ttatgctgta	atctctagta	ttaggtcata	cgtgggtata	gaacttttga	aatcatttga
12961	gtcttcatgt	tggtacttaa	gatccaatgc	tctcagttt	tggtggctgc	ttcttacaac
13021	tctatgatga	ttcatttttg	tgattacaga	actttccact	atgttatgtt	ttgggtcatg
13081	tggtttactt	tagttgtcta	gactaagctt	tggtcgtcga	agggtctctt	tctgttttgt
13141	caaccagatt	gcaaaacggg	aaaatgtgtt	aaagtaattg	gttcccagga	taaaaagtcc
13201	cgccttcgag	ttagttctat	ggccttgat	gggagtga	gctgggtggg	aagaactaag
13261	caactagata	atctaacatt	ttcgcacact	gctatttgtt	tgaagggtat	gttgctcca
13321	gtaacctaa	aatatagttc	ctttctcctg	aagaatcact	tagaagttaa	aatactgaac
13381	atgtattttac	tggtattggg	atgattttct	cctctttcag	gtttgtggac	agggcaaaaa
13441	tttagcttta	tggaatcttc	ccgcctcaga	atgctgacaa	acaataccca	tccctgcaca
13501	tgtacaggat	gtgatgtttg	atgaaaagca	agtaagcaga	acaagattag	attgaaattt
13561	aagcagtgta	atttaaagcc	ttgggaggcc	gtgggtgtta	tttgtttatc	aaggcttaat
13621	aggttttagt	agtatctctt	tgaatcatcc	ctcttcaacg	aagtttatta	aaaaaattac
13681	tgacttttct	aaactagaaa	cagagggtcaa	cacctagtgt	taattcctgt	tcaaaagtgt
13741	aaacatcttt	ttaaaaatgt	tgtctttaac	aagtatcgca	ctatcagatt	ttgactgtag
13801	gagcagaacc	acttctaaga	cgtttcgact	taaatggagc	tttgctttct	caaattcact
13861	gtgctccttg	ttcagtattt	tccatttcct	tgcattccagc	aggagtatgt	ttccaatccc
13921	ttcttttctc	cccatttcct	atatatgtgt	ttgcttttct	cccatttcct	atgttactgc
13981	taatatctg	caaatatgtt	acccttctg	gactagaaac	tcaatgttta	accgttgatc
14041	caattctttt	gtatttgcag	gtagtgcgtg	tgagggtta	tgagggtatt	gttgatgtca
14101	tctctcaatt	tggaagccat	ctctgcacat	ttcgtagcag	ttcattgtaa	aactccttac
14161	agtttcttga	tttggctcga	taccctaaag	aggtaccgaa	atgtgtcagt	cagattattg
14221	cagagggttaa	gaaacctttt	ctttgttttt	cttaggatct	tgctcctgac	taaagaccaa
14281	ttagttgttt	ggttaagtgt	aattgggttg	tggttccgta	gttgctatca	aatttgaacc
14341	attaaaatga	cattggaccg	gttttagtctg	atcacgtaca	ttactgccc	ttagagtttc
14401	aattttttgc	gagtcctgtt	ttcttatgac	caccattgat	caaataattt	gatccataac
14461	taaatgattc	tcactcttcc	aagtaaaaat	ttaccatatt	atgcaattta	atacatggac
14521	aaaacatata	gtttccatat	ccacaaaaat	aagtcctcaa	aaacatcact	aatctaaacc
14581	tcaaaccatca	ctcctgataa	gagaaaaccaa	actccagcga	tcaacaaaac	caaagaagaa
14641	gaaaatgaag	aagctgtgga	agcaacaatc	tgaataggga	aattgcattg	tccctgagat
14701	atattctttg	tggtaatggt	ggccaaacct	tggaagatac	aagcatcctc	atcctgggtc
14761	tttacttgaa	aatacatgtt	aaacgcata	gaagcggtcc	catttgcatc	aagagtattg
14821	caagaagacc	cgtaaccaag	cgcagtgcaa	tccgaaaaag	tgaggcgta	atctatgtta
14881	gcagcgagtt	ttgtcaaatc	tttggcttct	gtattaaaca	tacaccactt	cttaggttga
14941	tacgtcacgt	tctcggcacc	aatcaacaac	ttgttctgtc	cttgacctga	tagactata
15001	ggaaactttg	gttgcccgtc	gaatttgaat	attccccagt	ggcggttcaaa	ctctcccggg
15061	gctatacttt	tcgcatcctc	atcaagtaat	ccaaacaagt	atacctcaat	gtaggttggt
15121	ctcagtggag	tgctctgttt	ttccccaaag	cttggtaaaa	gaccgttgta	gaatctgtaa
15181	gcactgcctg	cgaaaaatac	gttttcaaaa	catttcattg	catctaaaga	actggaaaat
15241	gttagatgtc	ttaccactat	ttgcgtgttt	gtcaccctct	gttggccaac	cgacttctcc
15301	aacgatgatt	ggcatatctc	catgaccaac	agctttcaaa	gccgatacca	aagtgtcgaa
15361	attggcatcg	aaaacgtttg	tgtaagcaat	accattgtca	tctactgggt	tagcaccatc
15421	aaagaaggcg	taattgagtg	gaaaatcatc	attcccatag	aggcttaaga	aagggtagat
15481	gttaatgggtg	atgggtgcac	tgttattgcc	aagaaagtca	acaatctgag	tcatttgacc
15541	gatgatatca	ggacggaatc	ttcctgcaga	tggcacaggg	ttgcttgatg	gtgaatcgta
15601	gacatccgag	tttaaaggga	ctgtcgcttt	gacggagctt	cctagtccag	cttcattcaa
15661	agcgttttgg	atgttttgaa	gtgcagggaa	tgtgaggttt	ataaatgata	cattgtatga

19981	aataatgatt	ctgaaagttt	cagaaaaatg	tctggaaata	tataactgtg	ttgggttttc
20041	atttgaaaat	tatgatctct	tactagtaat	aacgtcatgg	aaattgcaac	acagaaaaag
20101	acttataaag	ttttctgata	tttttcta	ttaggatttt	cttttttaaaa	aatacaaaaga
20161	aaaaccgact	atagaaatgt	tggtgtaaat	taaacaggag	gaagagtttt	atccaataat
20221	acagtataca	agtaacaaga	tgaagaacct	cggaacttg	atagctttga	gggttaacag
20281	tgaatcatat	tttttatata	accaagtc	aaactagaga	aacctataa	attgaacaaa
20341	cgaaaaaaga	caatctcact	tccatggtta	gtaatctttc	atttagaaa	atcttaaggg
20401	aaatatataa	ttgtatat	tctctctata	ccacacaatt	ccgatgaa	cacatactgg
20461	tatatgttac	atttgtttat	aatgcttggc	acgacggctt	gtgataaggc	tttatcgtct
20521	cataaaagga	aaacgtagtc	attcgatcat	ctcccaaagt	ttcaatcttc	tctactgtct
20581	tcgtttattt	caaatgatta	tgagttgatt	aattattcaa	aacacagaag	atctctctat
20641	acatatatat	atatatatat	atagaaacct	ttcaaaccat	ttcgcaaatt	ggttgtttct
20701	cactttctct	agcgtaaate	tcgatgagct	taagttaaaa	cttaccttca	gggtcatttg
20761	ttttgttatg	tgacaatctt	ctagattaat	attccactta	ctacttcttg	cttaaaatat
20821	ttagttacat	cacatgacca	tgtaattgaa	tttatcctct	ttataatata	aactacgaaa
20881	atctgaagaa	gaaaaaaatt	atcgaaaaga	gaatcataat	ctggtactag	caaaaataat
20941	ttggtagaag	atatatatat	atttttctat	atgtaaactt	caaaattaat	gcctaagata
21001	tgctaaaaat	ttgcgaagga	gtcaggggga	aagcttgaga	ggaccaatgc	atggcattgc
21061	ttttactgac	agtaaacagt	gtcacgctca	cgaccatttc	ttcccgttcc	atttggtttt
21121	atttattttca	aagtttaata	ttccttttgt	ataacattca	aatcttcaca	tgattgattg
21181	tgtgaaaacc	ccacagattt	tactacaata	gggggagttg	acttaaaata	gctattgatg
21241	tcgaaaaaat	gtatttttagt	tataaattat	actaaagaaa	atttttgatt	tgtctgttgt
21301	ttaagcatat	gtattgttaa	acttaaaaaa	atatgtattg	ttaatcttaa	aaatgtagga
21361	gtacacatca	aatactcgag	cataatcaaa	accgtattca	tagaccgatg	tagaagataa
21421	atagaagata	atgtgatttt	ttaaaatatc	gtatctccaa	atcaatcact	tagaagataa
21481	tgtaattctt	tatgtgctac	ataaataaat	atatatatat	atatatatat	atatcttgta
21541	tatatgtctt	gacaaaaaat	tgccagtc	aaaccatgac	tgaatcaaac	tataagtcgg
21601	attgaatcaa	actataagtc	ggatgagtat	taatttccat	tatgtttcta	tactttacaa
21661	accgaaaaat	agatattata	gataccaaaa	aagtagattt	gtgtatatta	ttagaagatt
21721	tggaatttca	tcattatcag	gatctaaagt	acttccctaa	ttaaatcatg	tcgggtgaaa
21781	aagctcaatg	aatgtttgaa	atttggaag	tttattaaat	tcggatcttt	tttttttgtt
21841	tgctgtccca	aacattttta	ttttattaca	aataatcaac	ttatccttac	tactaaatca
21901	tttcatatct	ttgataccaa	caaatcattt	catattctat	tttgatgttt	aagaaaaacac
21961	tatttaccag	ttacaaaaata	ttataaggat	tggtgtttag	aaaaaaaagt	acaagttgaa
22021	ttctttttgt	caaataataa	attgactttt	taatataata	ttgacttatt	gaacatgatt
22081	acagaattaa	tcactacaaa	aactttccaa	gtttataata	aatacatttc	aaagactatt
22141	agttcttctt	aaaatatctt	taaaagtgt	caaagactac	cacatataat	tcagaaaaag
22201	tagaagttga	tttctttttg	tcaaataaat	aattgactta	aaatagtttg	gaaagccatt
22261	gaacttgatt	atagaattga	taatgtacat	aaaaaaatc	caagtttata	ataaatatcat
22321	ttttcaaatg	ctatatcagt	tcttcttaaa	atatttcact	aaaaaaacac	tcaaatatag
22381	aataaattta	ttgaataaca	taccaactgt	aaaacagaat	ttgacaaaaa	aaaaaaaaaa
22441	atgaaatgaa	gatgaagaca	aaaataaatc	accagaggat	cttatgcaaa	aaaatatatg
22501	aatacacaa	aaaccatatt	gatattttta	aaataaaata	aaaacagaaa	aatatcccaa
22561	caccgctttt	caattaaaaa	tcttccgtca	ccattgttgt	catcttcttc	tctcgtgaat
22621	cctttttcct	ttcttctctt	tcttctcttc	agagaaaact	ttgtctctct	tcttataagg
22681	aaccagacac	gaatcccatt	cccaccgatt	tcttagcttc	ttccttcaat	ccgctctttc
22741	cctctccatt	agattctgtt	tctcttttca	atttcttctg	catgcttctc	gattctctct
22801	gacgcctctt	ttctcccagc	gctgtttcgt	caaacgcttt	tcgaaatggc	gattttggat
22861	tctgctggcg	ttactacggg	gacggagaa	gggtggcgag	agttcgtcga	tcttgatagg
22921	cttcgtcgac	ggaaatcgag	atcggattct	tctaacggac	ttcttctctc	tggttccgat
22981	aataattctc	cttcggatga	tggtggagct	cccgcgcagc	ttagggatcg	gattgattcc
23041	gttggttaacg	atgacgctca	gggaacagcc	aatttggccg	gagataataa	cggtgggtggc
23101	gataataacg	gtgggtggaag	aggcggcgga	gaaggaagag	gaaacgccga	tgctacgttt
23161	acgtatcgac	cgtcggttcc	agctcatcgg	agggcgagag	agagtcact	gtgctccgac
23221	gcaatcttca	aacagggtta	aaatctcaga	aatcttcgaa	tttggtgttt	gcttggttgtt
23281	ttatatggaa	ttgagtttgg	tgattgtttt	gcattgcaga	gccatgccgg	attattcaac
23341	ctctgtgtag	tagttcttat	tgctgtaaac	agtagactca	tcatcgaaaa	tcttatgaag
23401	gtttgtgtgt	acttgtttct	ccttttagga	attgaattgc	ttgaaaattt	atcagagacg
23461	aataactttg	ttgttgctat	cattcatgta	gtatggttgg	ttgatcagaa	cggatctctg
23521	gtttagttca	agatcgctgc	gagattggcc	gcttttcatg	tgttggtaaa	agaagatgtt
23581	ttttattttcc	agcaatgtta	cattgtttata	cgtataatga	tgagtttagt	tgcaagtttc
23641	ctcttttgatt	cttcttttct	gttgacgtat	atccctttcg	atcttttctt	tggtgcctt
23701	tacgggttgag	aaattggtac	ttcagaaata	catatcagaa	cctgtgagta	attactattc
23761	tccagccatt	actgtaattt	ttattgaaga	caagtttgta	tcatgaagaa	cttacaagtt
23821	ctgttttgaa	aatgctcaag	gttgctcatct	ttcttcatat	tattatcacc	atgacagagg
23881	ttttgtatcc	agtttacgtc	accctaaagg	gatactgttt	ttctggcttc	agtttgtgat
23941	actgttttta	agtttagttg	tctgaccggg	tgatcttgaa	aatggacagg	tgtgattctg
24001	cttttttatc	aggtgtcact	ttgatgctcc	tcacttgcat	tggtgggcta	aagttggttt
24061	cttatgctca	tactagctat	gacataagat	ccctagccaa	tgacagctgat	aaggtaaaaat
24121	acgaaaaaga	agcgtatgta	ttagtcactt	gcactgtgtt	actgttttaa	ccaaacactg
24181	ttatgaactt	taggccaatc	ctgaagtctc	ctactacgtt	agcttgaaga	gcttggcata

24241 tttcatgggtc gctccacacat tgtgttatca ggtaactgca aagtgcacat accattctta  
24301 tacttgcaag agtttcttgt ctaaactctg gatctttgct tttccccagc caagtatatcc  
24361 acgttctgca tgtatacgga aggggtgggt ggctcgtcaa tttgcaaaac tgggtcatatt  
24421 caccggattc atgggattta taatagaaca agtacgtttt cacatcttgc tttattagtt  
24481 ttccctgggtg aaaatcatca tccctgcgtt gtcaccactt gacttcatgt tcttttgtaa  
24541 cattttgggca gtatataaat cctattgtca ggaactcaaa gcatcctttg aaaggcgatc  
24601 ttctatatgc tattgaaaga gtgttgaaagc tttcagttcc aaatttabat gtgtggctct  
24661 gcatgttcta ctgcttcttc cacctttgggt atgctgtgat cccatctctt tcaaaataat  
24721 ttgcaaatc gaaaaaccga aaaaggctaa atctcatagc aatttgatat ttttagtttc  
24781 ttagagtcgg tgatgtaatt tcagttactg aacgcaaadc tcttgccaa aggttaaaca  
24841 tattggcaga gcttctctgc ttcggggatc gtgaattcta caaagattgg tggaaatgcaa  
24901 aaagtgtggg agatgtgagc tattttactc aaaagaaaac ttatgatttt taatgttgtc  
24961 gttgtttttg ggtcatctaa ctaaccaaat tcatgtattc actgtcttcc tttatcagta  
25021 ctggagaatg tggaaatagg tatggttctc ttctaaaca tcaccttctt ttgtacacaa  
25081 aatagaagaa gagagctaata taagatcttg ttttcttga cagcctgttc ataaatggat  
25141 gggttcgacat atatacttcc cgtgcttgcg cagcaagata ccaaagggtg gtgagatata  
25201 taccgatatg caattgtcga gattgttttc tgtgatataa atttaacctt ccacacactt  
25261 gtttttcaga cactcgccat tatcattgct tctcagttct ctgcagttct tcatgaggta  
25321 tacatacttt ctacattgcc ctgtctctag acgcatgaac acacgtagt gaaagaaatg  
25381 ctaatatcca aagcattgtt tttacttaac gatcttgtgt tacaatttcc cttttgacag  
25441 ctatgcatcg cagttccttg tctgtctctc aagctatggg cttttcttgg gattatgttt  
25501 caggttaaaa aattactaaa ctgctgcagt cgatttttac taaactctaa tctcatattc  
25561 tgaccaacca atttgtttga gtaggtgcct ttggtcttca tcacaaacta tctacaggaa  
25621 aggtttgggt caacggtatg ctctcaaaac ccgagaaaaat agaacgaata actctttctt  
25681 tcatagccta gccattttaa tcgcaatgct gaaacttaat aataaagggtg atctgttttg  
25741 gaatggggtc atattattag gtgggggaca tgatcttctg gttcattctc tgcattttcg  
25801 gacaaccgat gtgtgtgctt ctttattacc acgacctgat gaaccgaaaa ggcagatgt  
25861 catgaaacaa ctgttcaaaa aatgactttc ttcaaacatc tatggcctcg ttggatctcc  
25921 gttgatgttg tgggtgttct gatgctaaaa cgacaaatag tgttataacc attgaagaag  
25981 aaaagaaaat tagagttgtt gtatctgcaa aaattttggg agagacacgc gaaccctgtt  
26041 ggattttgtt atggtgtaaa gaaatttcaa tcaaaaaact gttgtaatata ttgttaccaa  
26101 aaagaaatgc ttttctggaa ctagtaaaaaa cttttgttaa taaggaaaaa atgtgataaa  
26161 ggaccaaaatc tagtaaaaaa ccttccaata aagttgattg gtggtcccg tttggggatg  
26221 tgcattggga ttgtatgaaa ctttccaata aagttgattg gtggtcccg tttggggatg  
26281 gccattattt atttatcttt ttttttagcgt atttatttat gtcgtatgta tccaagggga  
26341 gacaagactc taaattgcaa taagtgttga ggcccgaaca tcatcattga caatatcagt  
26401 taatacatata catatggcaa atggttagaga aaatgtcgat gtgcagcaaa cacttttacc  
26461 cattcgaatt atgttatgaa gctttctttt accttttcaa acacttagct cattagatgc  
26521 tatataaagt gataccttaa atgaatttaa tactgaaatc tagatttctga gaagaaaata  
26581 tgcaacataa ctcttaggat atggaatact aataatctaa tatgtattta ataggtggag  
26641 caagcaacga aataatcaac cttttctttg tgttatatta aacctcatcg gcaaattatt  
26701 tagctttaat agatatactt tatctttttt ttggtgcgaa tatagagata cttatcttaa  
26761 aggtccaagt ctttaacaat ttgcataaat taaattaaaa tatttcattg tacaagaaat  
26821 tcaaatgaaa ctcatagtgt ataaacattt agtcgagtta caaagaaaca aagttattta  
26881 tggttacttc tttcctttaa aggaaaagaa aatggttacg agaggaaacca cgtgaagatc  
26941 acgtagagag gttggtcaaa catagaattt cagttggaag taaatttaat ttttaacgct  
27001 ccaccgactc tccgagacag ctgctctgta ctgagcgct catgttgact tggcagttct  
27061 ttattaatat tgtcgacttt ttttttttgt ttgactataa aagcgatatt ttgtgtccta  
27121 tttttttttt tttttgacaa agtgcataac tatttacatt taagattaat ttttttatt  
27181 actcaaatta gtagttatat atttcaattt aattcaatct gaaattcatc tcaattttct  
27241 atccacgaaa gaaagacatg aaaatcaact gaagtattgt ttctgttttt atatactttt  
27301 taagaatttt ttatgttact agaagaaatt tagaatactg tatatatatt tggatgaaaa  
27361 tttaaaataa tcttttagaag gcgaaattag aaaataacta cacaataaca ttatatctct  
27421 aatttttttt tttatatagt ttccaaataa acaaaacaac aactgtatca cgttttgtaa  
27481 atttcattta cctaatacag acattcttaa atttccaaa tttaaggaaa gtatatgtag  
27541 tcaacaaaaa tgattatcta cttaacatgg tatggttctc ttctttaaga aatcagttta  
27601 tataactaaa ttttgcaaaa taatgagatc gcattattgt aaatacatat tgtcttttac  
27661 tattttttat ttatatattt aaaaatgtta ccagacaaag gtaacataca atttatttaa  
27721 aactcgcacg aaaaaactcc attctcttac aatactttta caaaataaaa atgtaacaaa  
27781 taacttccat acattggcta tataaactct ttttagaaaaa cttcaacata cttatgtat  
27841 ttatttgaag tatcaatata taaaattgaa ttcgtcgact ttattacaaa accctctcac aacaccatca  
27901 agaaacaaaa aaagaaatca ttcgtcgact ttattacaaa acagatccag atcgtacggg  
27961 cttcttcttc gtcttctgtc tgagtccaaa atggaagact tctgtctacc gaagatccga cgttccagat  
28021 gacgggagaa catcagacct tcaacaatac tctgtctacc aggtcttaca gtacttctta cacagattac  
28081 tcattcagtg gtaacggtat gcaagatctt aggtcttaca gatcatcttc atcgtcttct  
28141 ccgaccggga taccggaaga ccagaacccg aagaagagag ttgttagtta cagagcttat  
28201 tggggatttg tggatccaga tttacaaagg aggttctttc agaaaaagct tcaaatggat caaagataaa  
28261 actgttgaag gtaagcttaa aggttctttc attctgtcaa atgtgtttta cataagataa taaacttttg  
28321 tgcaacaaat tacttaatta attctgtcaa atgtgtttta tttgtgtatg tgtgtgtgtg  
28381 tcacttcgat cgggttgagtt tcaatctttt tttgtgtatg tgtgtgtgtg ttaatagatg  
28441 tgttcttagc ttatgagttg tgatcgatag tgattttgta acaacaatac ttaatagatg



```

32761 atttctcgat ttctcatgcg gaggagagtg aagtgcgatg atcaaaaacta gggccttttgc
32821 gaagaagagg ggaagaatcg agcagggtga aattggaagt gagtaggtga ggaaatttag
32881 ggagattcca tgggtgttttg gttaagcaac cagaagcaag caagatgaag aagactgcgc
32941 gtgtgtgtgt gtctcgctta gcctgttgct tttggctctt tctattcctt ttctcaattt
33001 attaaatatc tttattcggt gaaaaaaata tacagtatat ttcttagaaa ctttataata
33061 agaataagga ttttgttttt atattttcct aaattccata aatctcttga gtttacaaat
33121 tcataatatc ttttaacattt aaacaaataa ccgacttctt agtaatatat ataaatttat
33181 ttagggagtg ttattggttt gtgatgtatt cgaagtttca gtgacttaaa atgttatgaa
33241 gaatttgttg ttattcaata aagatttttc aattttttgt taaagtttgg tgttatagtg
33301 ttaagatttg taaaaagtaa tataaaatct tcataaatgt agagtatttg gattcatatt
33361 tttataaagt tattaaattt tttgtgtaat tcatttaaaa caaagaatct atgattatta
33421 taatgaatta aattattgtg ttattggttt atgactttct accctttttt ttataaaata
33481 aagttgtgaa aaatatcata cccttttttt ttacaaaata aagttgtgaa aaatatcata
33541 aaaataaaga aattgttttg gaaactttac aaaattatta aaaactaatc aacaagttta
33601 ttaatgacct taacccaaaa taggtacata taatattaga tagatcatgc aaatgcgttt
33661 ggtgggttatg ttacttttgt atcatttgat tggcatgatg ttctacattt ttttaccatt
33721 tctctattct taagtcgaaa tctttaggac caaataagca tttttcccta aactaacaaa
33781 gcagcatctc actagatggt ttccaatcag atttgcttac tatagtgtgc taatatgaaa
33841 tttgagtgac tcatgacaat gcatcaatcc agattaagca ttaacatgtg atcagctagc
33901 attcatgaga atatgagaat tgtcttcgag ctaaaagaaa gtagactctt tttcacttg
33961 agaaactcat gacatacttt ttttttatac acatttgttt tttgatgatt aggttaaact
34021 tttgtaataa ataaattttt agttaaataa ataaacttct atcatcaaaa gactcgtgag
34081 tttcttcta tgttgcaatg aaatggaaac gaaaacgca atacgaaacg tttcggaaac
34141 gagaaacgat tttttcttaa aattagggat tggaaacgat tttttcttaa aattagggat
34201 tggaaacgta tacacataca tacatatata tatatatata tatataacat taacataaaa
34261 ctatattaaa gaaaaaatgg tttaaacaat acaagtccaa aacatcaata ttataataat
34321 aaactaaaat tctaaaagat aaaaattaaa aggttcaagt tcaaatagtt aatttagacc
34381 aacattttca ttttcattag tttcatcaaa gataatataa agctcttttc aaatctgggt
34441 tatcaaaaga aatatcttaa aactcaaagt cttccaaaat tagttatatg tattaataat
34501 ttttaataat ttaattttag ttattaaatt ttcaaagtg aaaaaaattt gtcgtatgct
34561 tatttgggaa acatgagttt ctatctttaa aacgtaaagt tccatcgtgt tcccaacat
34621 aaatttttaa aaaacgcgtt tccgttacgt ttttcatgtt tccggagagt ttctgttttc
34681 gtaacgtttc ggaaatggga aatagacctc ggggcgagtt tccatgcaac ctagatttct
34741 tctaaccaat agaaatgctc ttaaaagcct ttttttaaaa aaaatcaaca gaaatgctct
34801 taaatgttaa aaacaaatat tcatgccaaa ttttgatgta aaaatttggt attttcgctt
34861 tagttgtgtc ttatttcggt ctggtcattt tctcaaagcc cttttagtta tttatatata
34921 tattctctgt ctggtatttg tccccaaaaa tctagggttt taaggtttct tatccttct
34981 ctctctcgc cagattcttt tcttgcgaag atgagcaacg acaaggacag ctgaacatg
35041 tccgatctct ccaccggtaa gatattataa tcttattgtc ctatagaatg aggcctggat
35101 tctcttttgt tctcttgatt tattgaaaaa agcttctctt gttttgtgtt ttgcttatag
35161 ctcttaacga ggagatcgt gccgggcttg ttaatgctct taaggtgagt tttttcttca
35221 cgatatgatt tgcgtatgac tatctggaga ttgggctatt catctttgta acttttaggg
35281 attgttttct ctctctgact agttttgaga aattgatttg attcttatgg cttagagaat
35341 ttgacttttg ttttgaggat tgtctatgca tcttaatttg gttgtttgaa agtttgtgac
35401 ttttctgat ttgatttacc tgtttgtcgc agaacaagt gcagaatttg gctggacaac
35461 actctgatgt ccttgaaaac ttgactccac cagtcaggaa gcgtgtcag gcttcaagag
35521 agattcaggt gagtaaatat tcagctttta gatgatcttg gattttgtat tggttttgaa
35581 ttagctggct gttcaggtct aatgagtttt tgggtggagc aatttatctg attccttttg
35641 tattttaacc tttgcagaac caatatgat agatggaagc aaaattcttt gaggagagag
35701 cagctcttga agctaagtat caaaagttat atcagccttt atataccaag gtttgaatac
35761 ggtctttgat tctgcgagat tcttatggtc tttagtttcc tattattaga atatctttga
35821 aacacatgat gacacctcaa ttgataaagg ttttaataaga cttctctctg ttgctaacta
35881 cgtgaatgat tagtcttcca tggttttcag ttctttctgc atcattctg atgttaactg
35941 atattgatgg tcttctgcct tctgcatcac agcgatatga gattgtgaat ggtgtggctg
36001 aagttgaagg tgcagctgaa gaagtaaaat ccgaacaagg agaagataaa tcagctgaag
36061 gtgtgtttat cgattctttt actgaaacat gtttattttt agtatcttat gatgatgatg
36121 gtgaactcat gattttttata tgatatgaaa ctgtctttct gcagagaaaag gattaccaga
36181 tttctggctt attgcattga agaacaatga aattactgct gaagaggttc gttattagaa
36241 taftcttctt tttggtttat aaaaaggcga tctctttat cattatgtgg tttttccac
36301 ggttttagat aactgagcga gatgaagggg ctctcaagta tctcaaagat atcaagtggg
36361 gtagggttga agaaccaaaa gggttcaagc ttgagttttt ctttgatcag aatccttact
36421 tcaagaacac tgtcttgacc aagacatatc acatgattga tgaagatgag cctatccttg
36481 agaaggccct cgggtaatgt tttgctctat caagtattta ctgtttatgt tctgaagaca
36541 ataagtcttt attgactgtc gtttactgct gttcaggacg gagattgagt ggtatcctgg
36601 aaagtgtttg acacagaaga ttctaaaaaa gaagccaaag aaaggatcca aaaacacaaa
36661 gccgatcact aagactgagg actgtgagag tttcttcaac ttttctcagc cacctcaagt
36721 tcttgacgat gatgaggatc ttgatgatga catggatttt ccatctccat aagcagttta
36781 gtttttagag tcagtaatta agattgtgtt ggattttaat ctgatcatct aatattcaca
36841 ggctgatgaa ctccaaggac aaatggagca tgattatgat atcgggtgtg tacctttcta
36901 tttcatattc agttctcttc acttagttca gttctaggat ctgagctctg ccactgttta
36961 tcctgtagtt caacaatcaa agagaaaatc atctcgcagt ctgtgtcatg gttcactggt

```



```

28501 agtaagtgat ttcttattca atttttgacg tatcaaaaaca agttacaatg ttttaaacad
28561 ttttaattaaa acggtgatta ttcatttctt tatcattgaa aacatccatg agttattatt
28621 atcatgtttt gtagtagttt cgtttcatat tatgctgtgc taaaaagata acttgtttta
28681 tgcagtgtgat tatcttgtaa aaacgattag ttcaggtagt ttaagtaatt acggcatgaa
28741 tgttaaagtt cttacgtttt taatttgcga tactcttctt tcagttttga tgaataattt
28801 ttcttatgaa ttatgaagtg gttttgtttc attgtgtgtc taaaagtgtt tacacgaaac
28861 tttgttaatt aagtgaatcc aacgttttta gcgttacatg attataataa cgtcttgatt
28921 taggcgatt ggtgtgtttg agaaactatt tttatttgta ttgattagta tatacataag
28981 aagggaactaa ttttttgtaa acaatgtatc actttgtaaa gactagatat gacttataaa
29041 atatgttgtt ttttttaata atgattgggtg agtttgatga aaatggactg tggaggaaaa
29101 caaatagcag tcaaacaccg attagaacac cacttacgta cgtcttctaa tcgatatttt
29161 actgctactt gtgttcctca agagtacttc aaactcacat cattatacaa ataaaattaa
29221 tatactctaa tcatacttct tgaattttta tgaccaaaac aacatacaaa atcacatatt
29281 ttagaattca aatattagca tttctcttaa tattattata catgataaca acaaaacact
29341 tttgaatttt tggggctttt ccatacacia ctctgcaaaa tcacgaaaaa atgacatata
29401 tagcaaaaa aaactccac aattctacat gatggattaa ggatagtcac gaaataaacc
29461 tacaagtgtt tgtatttgaa atatatatac actttttttt aacatagggtg tacgaaggaa
29521 ttcttaacta aattataaag ttgttttgaa ttatgctatg ttctacgct tatacgataa
29581 atgatatatt tcttgttgat tactgaattt tggtttgctt atgcgataaa atgaagtcag
29641 acgagtcctt ctttcttctt gttttttttt tttttaattg tatattacgt gactttgttg
29701 tttgtttgtt gattccaatt tattacatgt tatgtttggc tatgcaaaaa tcttaaagcc
29761 gaacttttga aaatgaagcc ctgcatatat agtaagagat tacgttgtaa ttataggttt
29821 tgaatacgta aaagtctctt cctaccgaca tagtgttcgt cttcatgttt ctagaagta
29881 caaattgagt aagtaatttt tgtcaaaaac atttgaccaa ttagacaaat aaaagtcac
29941 ataaattttt aagatatgaa acataaaaac agatattaac ttactcttat cccccactgg
30001 attttatatc attccgcaaa tattatcaga gaggaaggat cttaaaattc gaagagtttc
30061 tctatacaaa tggaggagac aaatgaatat ttgaagaatc atagtttaag cactttcggt
30121 tgggtataata aatctatatt atctctctct tatatatcat cagttctcta gctgtgaaaa
30181 tagcctttct tcagaagaat catagacaac gtcactgagt gagtacagggt tttgttttat
30241 agctgctttg tatgaatcag tcagagagct ttgctttggg gttctacatc gtctccattt
30301 gattgtggcg ttgctgtgtt ctcactgcct gcattgttat tgatccgtaa accatcaatg
30361 cctttgagtg cggtttcata gtttgctatg tcaactccgc cactaaatgg ttacacagctg
30421 ctctgttcca caactcgtct tgacgatgat ccctctgatc ggcccaagtt caaagatcca
30481 ggtagctgag aaatcaatgg atcacaatga acatcccgac tcaaaagaag gtacccgaga
30541 agaaaaaaa acagcatgac aagaaactta ccagcttgct ctttgcaatt gcagaatcat
30601 ttccattctt gtctttctgc ttcagggtaa aatccggatt tgcctctga tttgggacat
30661 ctatggcaga caaacacgcc cttagtatat aaatcaatgg acaactatgc ggctgcattt
30721 tcataaagaa ggataataag cattcaaaaa tatacctgga cgcttctcag aattgccaac
30781 cgcagggttt aatccagaac tagttccaac accaccatcc tgaaaaaaga aacatagaaa
30841 gatcttgcac tattcatctg tgtatgtata tgtatcatga acgggataga aactttaaag
30901 tatcaagagt gtacatgggg acgaggttga gggtttccag attgtgattg ctgatacttg
30961 tatactgtcc agtcaaacac aaaatcaaac tgaaaacctg ggacaaacaa aatactcatt
31021 ttgagtcaac agatgcttct gtatgaactt aatagcctga gaggaagaag agagatagag
31081 aaaatgaatg accttcccggt ataaaaaggt tgcggaatag tctcttcaaa tatgcatagt
31141 ctggcttatc atcaaacctt agtgagcggc agtaatggaa gtaagatgca actctgttg
31201 gatgacctct gcataacgtc tgcaaaagtt cagtactatc tatgataaac cgtctctacc
31261 tgatcataat cctttgaaac atggattcgt ataagtttct aaatgtaaac catataagaa
31321 gaacttactt cgatggaagt tgaaaccttc ttttcgctga tcttatcata cttctgtttc
31381 ttgtttccag ctttcagccc ttgccaaggg agactacaag agaataagat attagagatg
31441 gattggggaga caagaatcag aacagaaaga aaagtactt aatgactctc accctccctt
31501 gaggaataac atgaggatgt aaccaagcga ttctatatca tctctccgac tttgtcttac
31561 caagtgaaaa atgtcaaaaa gtaatcagtg aaagaggaga aaaagatccg cgattatcat
31621 gagggtttat agaaagtgtat ggagtatacc aatccctagg tgagtgttca agtggcata
31681 cctcggagtc ccaattagac ttttattctc cctgaatcat ttacacgtaa atccaaaaca
31741 ataaagaacc cgttgctcagg aaatctgaag ctgacataca atatgctgaa aatgctgaat
31801 ctgtgacatc caaaatcttg ttacctgtat gggatagtgc tatgagttga gctgtctcta
31861 tatttcttag ccaaaccata gtctatgatg tagacctaga gaaagagacc aaaatcagtc
31921 atacaatagc gagaccccaa aaacaaatcc aactgagaga atttgcaaaa gtaagaaaac
31981 ctgatttgcc cgcttcccta agccccagag aaaattatcc ggctttatat cacgatgaag
32041 atacgactta gaatgaatga actcagagac atttatctga cacaagataa acccaacagt
32101 gagagagtta cattaagaat gacaagacac caatccaaag cataattacc atttgatcag
32161 caagcataag aacagtcttc aagctaaact gcctcttgca ataactaaac aaatcttcaa
32221 gacttggaac aagcaaatcc atgactaaga cattgtagtc cccctcaaca ccataccatt
32281 tcatatttgg aatcccagct acattaacga acatacacia gaaatgaata accaaatgag
32341 aattcaataa caatagagat aaacatttca ttgcatactt ccacctgaa gaactctata
32401 tatctctgac tcatatgaca attgcggtat cgcagctctt acactttctt agatagaaa
32461 aaaaaaaact aagttttgat gatccaagtt gggactttga agatagaaa
32521 cttaatagca acttcttcgt tagtctgaac atcagttcct gcaatcaaaa cagtctcatc
32581 aaacgactat tgaagagatt ttatatgtat agagagggga ggaacgaacc gaggtaaatc
32641 tctccgaaag agccgctccc aattttgcgg ccaaggcgaa acttgttccc aacacgaggt
32701 tccataccca attgaaaaaa ccctaacttt gtggttttat ctatctacta cgtcagagcgg

```

37021	gaagctgttg	aggcagatga	ccttgatatt	gaggacgacg	atgatgagat	tgatgaagat
37081	gatgatgaag	aggacgagga	agatgatgag	gatgacgagg	aggaggatga	tgaggatgat
37141	gacgaggagg	aagaagcaga	tcaaggaaa	aagagcaaaa	agaaggtaaa	ttatgtgggt
37201	ttgttctact	taaaaccttc	ctacatagga	aactaaaacc	tctgaaactg	atttgggtgt
37261	ttatcttttg	ttttgttgca	gtcatcagct	gggcacaagg	tttaacaatc	aatcaatctc
37321	gatctttttt	ttttgttgat	aatgcaatgg	ctaacctgag	gtatcctttt	ttaatgaaca
37381	gaaggctgga	agaagtcaac	ttgcggaagg	tcaagcaggt	gagaggccac	cggaatgtaa
37441	gcagcagtga	agaagtgaag	aatcttggct	tagttatgat	gaagaagaag	agtgaagagt
37501	gtcttttgagc	cgaggttggt	tttctttaat	ttgcagagtc	atgggtccggt	ttattatata
37561	tcagtttttgg	gtgattgggt	tgctatttaa	aaaaaaaaaa	tgggttcttt	ggtttggttt
37621	gtgtctcttg	atttttcctt	ttgtaatgat	cttatgaatt	tgtttcgagt	tgatgtcggt
37681	ctctgggtcag	atttcgaatt	caattctatt	tatcctccct	cgtaaatgag	agaatttggt
37741	agacaatcta	gtttacttaa	gattgatcga	attttataaa	ccaacattac	caaaccgtca
37801	aataattaaa	accaatcaat	cttatttatc	ggtttgcata	acccatcaat	gagccggttt
37861	agacatcggt	ttgagtttct	ctgggaaaga	caaaagtcaa	aaacatctct	gccgacttgt
37921	aaaagaccga	tcaacagaaa	cccaaaaaaa	atagttgaaa	cggaggggaa	cgaaaaccta
37981	aaaccctaaa	aaactcttcc	tttttttttt	ccagtgaat	ttctcttctt	ctccgttttc
38041	atacaagtct	gacttctggt	tagttggaat	ttccagttt	ttgggttggt	tctgtatctg
38101	tggttttaaaa	aagtggagaa	gaagcttttt	agtgcctttc	tatggcgagg	attctccgaa
38161	acgtttatct	actgagaagc	tctctgtttt	cgtcagaggt	atgtttatcg	tgtttctcat
38221	ttgggtatta	cgagaaatta	aaaaaactct	gttactgtcg	ttttcattgc	ttatttgggt
38281	attcattcat	gagaaagaat	ccgaaatggt	gtctcttaat	ttgaattcat	tctttttttc
38341	tgggtaatgt	ttattgacaa	gggtttcatg	gggttttgca	gttacttaga	agaagtgtgg
38401	ttggaacatc	gtttcagctc	cgaggttttg	ctgcccagg	ttgtcttaag	cttttacctt
38461	ttgtttctat	gaatcgattt	tactcaaat	gggttttgat	tttgatgaat	atatgtattt
38521	ccacagatta	ctagagctaa	atccgtttgc	gtttctgatt	gcttagctca	ctgtgggtgt
38581	taaagttgag	gtttttgcca	ttaaactctg	aattgttttt	gttttgaaag	taaaaagaaa
38641	tccaagtcag	atggaaatgg	atcatctgaa	gaaggatatg	cgaaaaagga	gattgctctt
38701	cagcaagcac	ttgatcagat	taccagttca	tttgccaaag	gggtccataat	gtatctcggt
38761	cgtgctgttt	ctcctagaaa	tgccccggtt	ttctctaccg	gatcttttgc	ccttgatgta
38821	gctttgggag	ttggtggcct	tcctaaggta	tatatactta	tctcttttgg	tgatattatt
38881	ctttcccaa	tatgtgtgtt	ttagagtttt	tataggttgg	ttctattagc	taagttaatt
38941	gaggtttatg	tataagtctg	tattctcttt	gaagaatcta	ataatattgg	tctccccatt
39001	gtgaattcca	tagggggcgtg	ttgtggagat	atatggctct	gaagcatcgg	gaaagacaac
39061	acttgctctt	catgttattg	cagaagcaca	aaaacaagga	ggttattcct	tgtttctttt
39121	aactcctcgg	ctatgttctt	tacagagcca	tacgtcgtat	cctaaagaag	ttttttgcat
39181	acaggaacct	gtgtctttgt	agatgctgag	catgctcttg	attcgctact	tgctaaggca
39241	atcgggtgtaa	atacagaaaa	tctgcttcta	tcacagcctg	attgtggcga	acaggccctt
39301	agtcttgtgg	atactttaat	ccgaagtggg	tcagttgatg	ttattgtagt	tgacagtgtg
39361	agtaaggtga	ttatatggg	atggataatg	ggttgatgct	tttgctattg	gatctatatt
39421	tcgctctctc	atgttttcat	gtgattttgt	tttacaggtg	gctgctcttg	taacctaaag
39481	agaacttgag	ggcgagatgg	gtgatgcaca	tatggctatg	caagccagat	tgatgacca
39541	agctttgcgt	aaattgagcc	attctttatc	gttatcgcaa	acacttctga	tctttataaa
39601	tcaggtaaga	gaacgttact	agctgagatt	tgtattcgaa	atgtaaagtc	tcttatgcaa
39661	atgtatcttt	acttccccat	gtttcatatt	caggtgagat	caaaactatc	tacgtttgga
39721	ggattttggag	gtccaacaga	agttacttgt	gggtgaaatg	ctttgaagtt	ttatgcttct
39781	atgctgttga	atatcaagcg	aattggactt	atcaagaaag	gcgaagaggt	aaacttccga
39841	aaccgcgtta	cacatttttg	gttccgaagg	cttatcctac	gacctgtctt	atgttcgtga
39901	tgtgtgttta	actgatcact	taacctttct	attgtctcct	tagttcttca	gaatgattaa
39961	atgcttgtgt	ttgaaacctg	agattgtatg	cttgtgcaga	caacgggaag	tcaagctctc
40021	gtgaagatag	tgaagaacaa	actcgctccg	ccgttttagaa	ctgctcagtt	tgagcttgaa
40081	ttcgggcaaag	gaatctgcaa	gatcacggag	ataatcgacc	tgagcataaa	gcacaagttc
40141	atcgcgaaaa	acggaacatt	ctacaatctc	aacggtaaaa	actaccatgg	aaaagaggct
40201	ttaaagagat	tccttaaaca	gaacgaatct	gatcaagaag	agctcatgaa	gaagctccaa
40261	gacaagctca	tcgccgatga	agctcgcgat	aaggaaactg	aatctgaatc	tgaggaagaa
40321	gattccctga	gagttgtggt	ttcacctgac	aacacagatg	atgaatcacc	agctcttggt
40381	gttggagctg	ctgcagtggt	tgttgaggca	gcatgattag	cgacctccgg	ttagtataaa
40441	tattcttctc	ttggcctaga	gttttccggt	ttaacgcggt	ttggattcgg	tttcttctc
40501	ctcatgtaat	ttatgtgcta	gttaaatac	atttacatat	aaccgttggt	gtgggtgaga
40561	aaattttgta	gtttttatgg	ggaaatttaa	tgtaacgaa	aagcagaata	tttaaatggt
40621	attgatcaat	tttactttca	ccaaaatgct	attattatat	agtaactata	ttatagtttt
40681	agatattaga	cctcacaaat	gacatatcac	attaagttaa	tcaaaagatg	tcttgccctg
40741	gagataacag	tttcttggga	tcaaatagat	ctttcctctt	cgaaaaatca	tcccattttg
40801	atccaaaatg	ctcaatccaa	tcttctttac	tagtataatg	cattagatat	tgcttaatct
40861	taatacctga	atccttgcaa	aacctaat	tcttctcggt	aacgctctcc	acttctggaa
40921	gaccccttgg	ggtagcggat	tgtagtagtc	cgataatata	tataacatct	tcatcgatct
40981	ctgggtatcat	cgccgacata	cgattgtccc	atctaacaat	aacagcataa	tcagttcata
41041	gaaattttatt	tacttttggt	cttgaagata	aatatgaatg	aagagaagta	tgtacttatt
41101	ccggtttggt	ggatagagaa	gagcgagtcc	cgaagctgat	ttttgcttaa	gaagaatgtc
41161	tttgacaaca	ccgttatgaa	aatcgagaat	ccgagattta	ggaacgtaga	ggtaaagcca
41221	aggatgagga	agttcccata	atcccaaaga	tctgagttta	ttttcttcga	catgtacacg

41281	gttcaagaaa	tcgaagtagg	ccacgtcgtg	cattgatatg	aaccgggga	agtaacttaa
41341	tgttttcgtt	aatgtgtcaa	taacctacac	aatagtagat	tatgagaagc	ataagatgca
41401	aaaaaaaaatg	taacgatgat	gatgatgaat	atcttaatgt	gggtgcaatga	tgaatcattt
41461	aaatatctta	tggtatgata	aaaacgatga	tgaaaatgta	aatgtgtagt	accttgctga
41521	tgtgggggag	attgggatca	tcataatact	tggtacttct	aagaacatag	atgataccgt
41581	gttgcttgac	tagatcacgc	acttttagatt	gatctgaagg	tggggaaaaa	agggtgtcaa
41641	cgacaccggt	tgatagaaat	atttgacctt	ctaaatagtc	gactccaata	tggtttgcca
41701	ttgatatacaa	acgttcttgg	tcctttgttaa	aagttgtgaa	atcactgtag	agcatccgaa
41761	accatttggc	ctacacaatt	tcatttagag	aaaacaatcc	cattttaattt	atacacacat
41821	atataaagaa	accatattac	agtatagtat	ttcaatatatt	tcttgaaaat	aattatataa
41881	aatttgtgaa	tagatgaaac	taaaacagaa	cgtcacgcga	tattatgacc	aaatctgttt
41941	tttaaaatga	taaccaaatt	gtaacgttta	tcggttgtaa	gatttctaatt	agtcagaaga
42001	ctatatctta	aactttggag	tatatttgag	tagttgacta	atgaaagagc	ttgactctaa
42061	taaaacaagga	ttatatattg	acatacaaaa	gatttgcacg	cataagacaa	gattagacaa
42121	caagtcctag	gattacaagg	actatacttt	ataaaaataat	ccatatcttt	tttttattca
42181	atgcatagat	tttgttcaaa	ttaatgcaac	ttaagcattt	gacctaaatt	ccaaacttgt
42241	acctaaacttt	agtttttagga	taccatctaa	gaaagcttca	tgttttgtga	tccatacgtc
42301	tctaactctc	ttagtgtaat	atcataaaat	gttttaatca	gacattatcc	gattagtaaa
42361	ataattattt	ccgttaggtt	atatttttatt	ttatctcaat	aaaaatcact	aaataatcaa
42421	gctctgacat	aagaatcatt	ttctattgac	taaaaaaatt	tagaaattga	aacggaatta
42481	ctaatgaaaa	atgaatgctg	actaaaacaa	atggtaaaat	atgatgatac	gtaccggttt
42541	aggtgcatgg	tccaaaacaa	ttctggctct	cgttataatt	ccaaattgac	ccaaaccttc
42601	taacactcca	tagaacaatt	ctgggttttag	ctgtcgcgag	catgtcaaca	tttcaccttt
42661	ccctgaaatt	atttattcaa	aatccatatt	aacaataaca	atttaaaaag	tgtaaatcca
42721	aagaaaaatt	ctgaaatggt	gctatttttaa	aaatatacca	cctctttatt	cgaagaaaaa
42781	taaatgatata	atttgaaata	ctacacaaat	ctaagaaaac	cgaaaagctc	catatgcattg
42841	cgctatgcat	gtacattaca	aaaatcccaa	attaacacgc	taccaattgt	gaaaaatttgt
42901	ctttaattat	ttatataatgt	acgacaataa	atgtgtgtat	atgtaattac	tttattttctc
42961	atgaacaaca	ttcaatatgt	acaaaaacaa	tatagatgcg	ttacgcattt	tgtgtttcga
43021	agacttacgt	gtggtgcaag	aagtttagca	caactataaa	acataaaacc	tatgggtacaa
43081	aaaatgaaaa	atactataaa	acaaaaacag	tttttgttgt	tgtatgtaca	tcaaagttaa
43141	gaagatgctg	accagtaata	acgtccaatt	caaggacgtt	actaacaaga	ggaccgtttc
43201	gaaacacttg	accaccaatt	ccaccattcg	acaacgttcc	tccgacggtt	atatgcaaat
43261	aatccgtcca	agaaaccggc	gacacccctt	tctccgccgt	cttcttaagc	acatccaccc
43321	ataacgtccc	ggccgccacg	tcagcgtact	tcttgtcttt	tgaaccacc	acgtcagtga
43381	tacacgtcat	gttgacgatt	actccgccgg	agaccgagcg	ttggccggtt	aaggagtggc
43441	cttggccacg	agccgctact	tggaaatgtac	tttttccggt	tgcggcggtat	tggaggagac
43501	gagagatatc	agcgggtggag	gaggggcaga	ttacgccgcc	gggggtcacg	gtgggttatgt
43561	ttccgaagtc	atgagaggct	gcggagatga	tgggaaggatc	ggtagagagg	gtgaggttaa
43621	gggatttagg	taaatcaatt	ttaataccgt	ttgatgattt	ggtgatcatt	aaaaccgtga
43681	ttaaagtgat	cattaaacga	agattagcca	tttgtttatg	tttctctctc	tctctgattt
43741	gatttttgaa	gaagagaatt	acaacgaaag	aattagatat	cgctataatt	tgttacttat
43801	atagacaccg	taggaatgta	tgtataacat	aaaatatgga	atcctaccaa	aaaaaaaaag
43861	acaataaata	ttgaagtcaa	aatgaaatct	caaaaatttc	ctaaaaatga	aaaaaggaaa
43921	tcttactgtt	ttgtcaactt	gttgaattaa	tacttcaagg	cacatcctca	tatttagata
43981	tataataaat	cttgttgact	atagttagcc	atatctatcc	ataaatatct	ttttagaatt
44041	tgaatttagt	attaaccaac	aaaaaaattt	acattaattt	tcttggaata	aaatcgagta
44101	aagctttcac	taaggtttgt	tgagaaaaat	attaaagagt	gaaactgaat	taaaaaaaac
44161	ttatgtttat	gtatgtgcat	ttccatgaac	cagcatgtca	taacataaaa	gcctagaggg
44221	gagggtcgag	gggagagtcg	ggtagttcaa	ggagttgatt	tggttgcaaa	tgatttcggtg
44281	aatatgccgt	tgaaagtaaa	ctgccgattg	gagaggcattg	cattacgttc	ttgccaaact
44341	gactaaagat	catatacgta	ggctcgtggtt	tattttttcgc	ggaacccaaa	cttttgggat
44401	tttgattcaa	agacatcata	acagataaga	tgatcgagat	gtagtaagac	aaggaaatgtt
44461	ttacatggat	gtaggcgtgt	tctgactcat	gtaagtactt	agaacgagga	tttacggaaa
44521	caaccctttt	atgttatgaa	ttttacgttc	tcgaaaattg	tgaatttttt	ttagagtagc
44581	ctctctaata	ggtttatata	tgaatttttag	gaggattatt	tggaaaattc	cttccgctta
44641	atgataaaaat	tcattcgttt	taagtttttaa	ccttatgttt	ctcctctttc	acggtctatgc
44701	tctcgactga	tcttttcata	actataatata	tgacatagtt	tgagtatcta	agtactgtga
44761	agtgaagatt	agtttttttat	aaaatatttt	aaaagtctat	gccaccattt	ctctgacaag
44821	ttatatctat	acgatttctc	tctctctcaa	ttgataaaac	aagcaagaaa	taaaataaaat
44881	tgtttttgatt	atagtagaac	actatagttt	gataaacgaa	gaacgagaaa	aaaatctaca
44941	aaattgtgat	tatagtagaa	cactatagta	gtcagagttat	ttatggacac	caaaaataac
45001	cactagaact	tttaattctg	gaacaaaagt	gggaatatct	tggtagatat	ggcaccgata
45061	atcacaatga	tatgatgttt	cgctaaagtc	atgacagggtg	gcgaaatggt	ctttgttttag
45121	ttatagttga	ttacttgatc	ttgttttgca	tagagataat	tgcaattttt	ctttgttctt
45181	tactagtata	actaccttaa	actattgaag	agattttgaag	aatatcagaa	ctgttgtcat
45241	atgtgtttgc	agtttctgtt	ctataacttc	tgttaaaaata	atcactttct	gacattgttg
45301	cagatattgg	ttgaggaatg	ctctgagtg	tcagagaacc	gtataggact	taatcgagc
45361	gggtggagct	ggatgcgtct	taaagataac	gaattagttt	ctggttcttg	tgatcagttg
45421	tgtttaattg	atattcatgg	gtgatagatt	taatcactca	taatcttatg	ttacaatcct
45481	ttgtgctatt	ttagttatgt	tcattagagt	atttatgata	agaactcact	agctgcgagg

```

45541 aggagatgga acttgccaga gaaccactca gattcaagag gaggacaaca aagatgaatg
45601 gacatgggtt tgggtacttat ttggaagaat actcctacta ggttactggt tgagatgatg
45661 gataagcttc tgaatgagag aatgacgaag aggagtggat gtgtcatgta aaccaaatac
45721 tgagctacta agaactgatt tcaagaatgt gtttagatgt tgatttggaa tcatagcaaa
45781 accatatgtg atacgttttg gatcagaaaa aaaaaacgat ccaaggcagc caattgttgc
45841 gggcaggccc gtctgccatc tttagataat tagatgtaac taaaatatga ttattattat
45901 taaataagaa tggattttgt caatttgttt ttgtctaatt acttatctga ttagaaatag
45961 tttgatttca atgattggat gattagtcca tctgtccata tactatgaaa actatgatat
46021 ataccacatt ttcacattct accacttctt tcatcctttt tttttaaatc taccattatt
46081 cccaaaccaa aaaactaatt tacactcaaa tctatctttt tagttttatac tgttactagc
46141 aaaagtccaa cgcgttgcat ggtgtaatta atatacttaa atgtataata aaatgtaact
46201 aatatatac tgaacattt tttgttactt ttaaaatata ataattatat ttatatttta
46261 tattgacta atttgttgtt agtggtagaa gctttttgta cactacaaaa ctaaacattt
46321 gtaatatgtg gtttttaaat agatgatgtt ttcaactttc aataaataat tgattaactt
46381 tttagtattg gttttaacta ttatactat gttaaattatt ttattattgg ttgaattgta
46441 ttcaaacatc taataaatga tgatttagtt tagaaaataa aaaaaataa attagattaa
46501 tgtgcaatag aaataaagaa ggatcatcat gtctacttaa gtgaaaaaac attagatcca
46561 aaatacgtt tgggttaaagt gcattttatt aaccgtgtga atgaagacca tatgtggcaa
46621 cactacatga agcttgtttt attgaaatat aggtttttta taggctttta ctgcacggtg
46681 gccattcaca gaatcctatt agacaaaaaa gaatagtcac aaatcctaaa tatctttagt
46741 ttgatatttt ggaaacattg acaagtgcac caataaaaac ttgcactaaa attctagtga
46801 aagtcggatt agataccaaa tattggaaag aaaaaatgat gcatgctaga gacacacact
46861 taaaaatttt gttaaagtga acggattaaa gaaccaaatt aactaaacca tttatcaagt
46921 taactaattt accaccataa ttaattttta atattaaaac taataatac aatgatcatt
46981 ttaaattgaa ttttaatttt tatataaata gaaaaccatt ttccaataat atttctataa
47041 tatattttat catttaataa taatatatta gcatcaaat atattaataa tgcaccacaa
47101 cctaatgacc caaaaacaaa gacgaccata tagaatttaa acttatgaaa atcttaataa
47161 tagttatgta gctaaaacta tacacgagat taatatatat tgtgggtggtc ttgtgacaat
47221 gaagacaaca acgacgcatt ctacaaaaat aagtcaaaat ttctattgga atcatgaacc
47281 cactcaacca ctgttttctt acacaattgg ttttttagcg taaaaaacac cgtgtaatag
47341 taaaatcttg aaattgagtg tgagaaatac ataacgtcgc taatagatac atgaactgtt
47401 cgactactgt ttttatacac aattagtctt tagcggaaaa tatcgtgtaa tgataaaaaa
47461 tttgaaattg agtatgagaa ataataatct tgaaattata aactttttta ccaatgaaaa
47521 tcataccgag agtttaaat tatatctaga aaataagcat ggccattaga caccacaaaa
47581 actcttaatt tactcactta gaagttaaaa cacattttcc ccattgtttt tgatggtagt
47641 cattcaataa aacgttagta attgaataaa taaaaaatta tggtaatgtg gatatatgtt
47701 agttatatgc aattatgcatt ttaataatag gttttaaaat gaatgagtag tcaacgacgt
47761 cgctatcgat ttttttttgt gattagtatt tgatttcctt tataattttc tttttttttt
47821 tttattttct ttataattac gatcatatgt ttataaaaat aatattttta atgattttat
47881 gattagacct tcactattgg gtcttcatat acaactctta caatttcatt tgtcattttt
47941 actcatcttt gagaaacatc agaatcgtat tttacttcat ctgactctgt ttgttttacc
48001 gcttatgttg aaccgatgaa attattgttt ttgttttcat gtttttattt atttatttta
48061 catgactaat tttatgattt aaatccatga ttataaacct taaattgttt tatttttagtg
48121 gatatttata atttagggaa aattagtagg atttcctatt atttctagta acgtgaagtc
48181 attgtagaaa acaaatagtc aaagaagtta tgggtccaat ataaatcatt ataaatttag
48241 atttttttct aattttttat ttaatttttt aaattaaaat ttataataa aaatgaaatt
48301 ttaaattttt agacaccaat aactaaaatg ttaataaaaat ttttttttat ccaaaattat
48361 aaaatttaata tcataacatc taaaaattaa tagttaattg agatgataag ggtaaaattg
48421 aaatttttatt tttgttatga gatcgtgagt atgaaaataa cgacatttaa ttgcttgggt
48481 attaaattgc acagttttat ttttcatgt atttaccttc caacacaaaa caattaggta
48541 ttaaacgtcc aaacaaagaa tatttttaggt ttgttttttg atttttttcc tttttattgt
48601 ttatagttga ataattttat tcccgtgtaa agttctcaaa cgatacgttt cttctttatt
48661 atataagata ttatttcttt attttatttt ttgttttctc tctctttctt cttctttttt
48721 tttctctctt aaatatatta ttctctctt tcttcccttc ccttgaattt taccacaata
48781 ataaattttc aaaagtcaaa tagacgatct aaacttgttc tcatccact attattcatt
48841 cgattgtaat ttgttcccga tcaaattcta ttactcattc aatatgaact tgtccccggt
48901 ataattatat tactcattcg atttgaactt gtccccgata aaggatttag atctggacat
48961 gccatgcaag atgggttaata taattagatc gggacaaggt ctattacttt ttctatagtc
49021 ttttttaaaa ttaacactct gttaaattta gggacaactg gctttgatga ttgtccttcg
49081 tcatatgagg tgggtttatc cgctagaaat cttaacttac ttttcgggga tttcattcac
49141 tatatttcatt aattattgtc tacggtaaat gtttaactaca tttaatgcaa cattttgttc
49201 ttattctccg ccataaagct gtttttctgg cacaagttac catttttctc taacacaaac
49261 tttcacactt ttaagggaca aaattgtcaa taaaatacta taataccata aatattggtg
49321 ttttagaaag gtcaatacaa ttgtggtaga ttgggaaaga atcttttcaa aagttgtaaa
49381 ctagtcaata ttcccaaaat ttagggtaat ataatttag tttttttcaa gtcaaatgta
49441 atctcaattt acttccaaaa caaaacaaaa taaattaatg tgttgaattt ttaatgcaca
49501 tcatatattt ggatctctaa ttaattttct tgactgcatt ctgcaaacat tatctataaa
49561 tacttatttg atagtgaat tttattctat ccttaaatca ctcaatcaat ttttatttct
49621 aaaagaacaa attattgtta tggaaaccact aggagatcgt cgctccttgc gcgtatgtat
49681 caccaaaaaa agaaattgtc ctagattctg cgaatatgcg gaatactttc catatgagtt
49741 qtatactttt tttaaaaact ttttttctt actattatct tctgctaaac attttacttg

```

```

49801 atattaatta attttaattt tcaggcgaag tcattatgaa agtactaatg aattgtttgg
49861 cacaccaaag atcattaaga tgatgaggca tgctcctgaa gaaaaaaaac aaatgcttgc
49921 aacttccata atcatggaag gtaatgcttg gacaaatgat cctgtaagtg gtggatttgg
49981 tatggtgcaa aagatcatgt ggaagattat gttacacaaa gcctaccttc atgaactcga
50041 ggagaagatt aaggaagaga aggaaaaaat cgagcttcat ctttaagtga tacattgtag
50101 catatattat ttatgcatca aataatgatt atttaataa aaatcatatt taagtgaat
50161 attaaattac cgtcgagcaa gttaatacct tatatttctt attcataatt cttgtgatta
50221 aattgtttgt gtatttagat gatcataaat gaaactttga aatctagaga aacatattgt
50281 acatataata aagtaaaagt ataaaaaatt gataatttta aaaatacgag aacttctcgt
50341 ttcttgttct aagtaaaaata acattccaat atagaattag ggttttagtg cctttggaaa
50401 attgctttgt aaaaagctag atattatggt tcgttttagta tatctttgac tcatgatatt
50461 tgacatttga ttaaaaatca tgtttgcctt acatgcataa tatcttaaat gcttgtatca
50521 catgtctcta actaccacga aaaactttgt gtataaaaaa ttcaaatctt taatttttgg
50581 acgacaaaatt cattagcgtt actaatttca ttgtacctat cagttttatg acgataataa
50641 tacaatttat ttctgtcccaa aaaaaataat atttatttat taagttaata ggaaaatagg
50701 tgtaaatata caaatatgta tatgtaatat gcaatttagag aatatatgaa acaccactat
50761 actatcaaca ttaatattag taagatgaga aacataaatt agatgatttg agtcaatatt
50821 tttccctagt aatcttagta atagtttttt ttttctttta tttgcatagt tcatatacat
50881 atttgtatac atatatgttg tgcagatcta aaatttttgt ttaataccta ttttgcacct
50941 ctaccaaaata aatttgttta gaatttttaa aaattcataa tcacaacaaa atttctaaca
51001 ctattttaatt tgttgatcat aaaaaatctg aatttttatg cacaaaacgt ttcttaatag
51061 tcaaaaactag tgatataaga agtgcaatgc caagaattat agagatttga agcgaaaattt
51121 taaaatgaca tgattataca taacttttaa gttctaaact ataatacata taaaaaaaat
51181 taagatatta aatagagcga acagacgttt aggttttttt cggatttacc atattcatca
51241 gattaatcca tgctggattg agaaaaagcc tcaagcttgg aacaactaca tgtaaatcat
51301 aacctcgaga tattttaagg ctttcaatgt ccāaaagaaa gttatttgggt gtcgaatcca
51361 tgagtgaaaa ttgagctttt cacatagaac aaagagggat atgaaatagt acaattttat
51421 gttttaaata aagaagataa aatgaaaggt aaattttata ggtttagcca gcagagaaca
51481 aatgaaatga atgttttaag ttttagacacc aagaacaaac ttagggttca agtctaaagt
51541 tagtcatata cttctgaaca atattggact ttattgttca agtattaaac aaacattgaa
51601 atcttcttaa atcacttttc aaaatgcaac gaggatctat tctttttttt tttttcagaa
51661 gaagctacaa gaaagaacaa caagtttgga tatgaattat aaagttagta atccatgtaa
51721 gagacgctta tattaattaa tttatactgt taattatttc atccgtaatt tttattatta
51781 ttcaataaag tatttatttt actgaattaa aaaataataa ttatgggtgt ttggatattc
51841 ggtcgggtat tttgggtttg agtttttcgg gtttagaatt ttaagatccg ttcgagtaat
51901 tcaagattcc gggtcgggtt ggttcgggtt tcgggtttag ttatatcгаа aaatatcaaa
51961 attttgtgtc cgaatctatt aaattatttg aaaatttcaa aaattcccta aacaacccga
52021 gtagttttgc ttgaatatat ttaaaaatac ataaaagtaa ctaaaatatc cgaaaaatca
52081 taatattgtc tatatgtaaa tataaatata aatatttttag ttatatttat atttaagata
52141 tgtttgggta ccagttcggg ttcgggtttt tcagggtttg aagtttagat tcagtcggat
52201 atttgaaaaa ttcaggttcg aatttagatt gggatttttg gattgggatt tttggatcgg
52261 gttcagatcg gttttttcgg attcaggtat tatgtccaaa gtaaaaatgg taaacggtgg
52321 tgtttttggt tgattttgta aagcagaaga gagagagata tagagacact gaagcagaag
52381 accaaaaaag aaaaattaaa aagagagaga ggaaaatgga gagcgacgaa gcagcagcag
52441 tgtctcctca agcaacgaca ccgagcggag gaaccggagc ttctgggccc aagaagagag
52501 gtcggaaacc taaaaccaag gaagattctc agacgccgct gtctcagcaa cagagcgatg
52561 ttaaaatgaa agaaagtggg aagaaaacgc agcagtcgcc gagtgttgac gagaagtact
52621 ctcagtggaag aggtctcgtc cccattctct acgactggct cgctaaccat aacctcgtct
52681 ggccttcaact ctcttgacgg tctccccctt tctcctttcc tctccttcta gggtttcggt
52741 tcgtaactcg ttcttagctt tgaacattct catgttttga atgaatttga taaaatctta
52801 cacatacatt ttctcgattt ctgggtttta gtgagattgt tgcgattggt ctagttaggg
52861 ttttggatgt ggctctgtct tcataccttg atatatctga tgttctattc atgaattggt
52921 actattgatt acctgttggt ttactaatga ctaagaggaa ttttcagttt ctctgagtgt
52981 ttatatctga tgaagtcttt agttgtttgt ctaagagttt ccatttggtg aattgttggt
53041 tgattttttt atagatgggg tccgcagctt gagcaagcaa cctacaagaa tcgccagcgt
53101 ctgtacctct cagagcaagt aagtttttag ctttctcttg tatcttgttg tctcatcttc
53161 tttataact tctcatcgta ttattttag ttttcttggt tgtgtcacca gactgatgga
53221 agtgtgcccc atactttggt catagcaaat tgtgaagttg ttaagccaag ggttgctgca
53281 gcagagcaca tttctcaggt attatgtggt ttaatactaa gcttgtgtcc tttccatatt
53341 ctactccaca ctacaattgg tttcatgttt gacacttata tactatcttc tgaaaatgtg
53401 ttctcagttc aatgaagaag cacgttctcc atttgtgaag aagtacaaga ccatcattca
53461 ccctggagag gtgtgaattc tgcccactct tgagatattt ctgtattgac attgttcttt
53521 ttggttctat ttggtttggt aattgtatct gcatccctgt ttcattctgc catgaagtta
53581 ttctgttggc acgttggtga aagtaaaatt tgatgtgtat tcattactaa ttgcaattg
53641 cagggttaaca gaatcagggg actcccacag aatagtaaga ttgttgctac tcacaccgac
53701 agtcctgatg tgagtgtgtc ttctattttg ttatgggtcat agcaacttga aatatgtcgg
53761 tttcatattt ctgtatttgg cagtcaaaga gcatcctttg ttcggacata tgtccagttt
53821 cagagttatc taaatacaat atgttgattt caggttctca tttgggatgt tgaaacccaa
53881 ccaaaccgtc atgctgtgct tggagctgca aattcccgtc cagatttggt atgtccactt
53941 ctgagaatgt tgttttatgc tttattcttg tttgtttctc atcattggaa gtgataaatc
54001 tctttgatatt cttcttaaat agtgcttctt gtttgcata tctgaatgaa ccatttttca

```

```

58321 aacctaattg tgtgaattgg atttattttt gtttggtttt gaattataat agaagaaagc
58381 agaggcgtct tcttccgaga taccatcgat accaactagg gtttggcaac cagggtgtga
58441 tacccttgaa gatggagaaag aacttcagtg tgacccttct gcttataatt ctctccatgg
58501 cttccatggt gggtggccct gtctgaggta atatcatttc tactcctata catgtgttca
58561 tgaagctgaa agcgtgagac ttaagggata attcaaata gaagcttctt ctttgtttgt
58621 caacagtgta aagaaatgtg tttgcaagtt tacgattcat tttcaattat ttttctcagc
58681 ttgtgaatga ttacagtga agttttgggt atttgatggc ttttctcagc
58741 tttgacattt taggtgataa gttgggtttg aaccgaactg agtttcttca cacactttat
58801 atggtggctg ggactcaggt aagttttgta ctttttataat cttcatatgt atttgtgctt
58861 ttagaatcct gtgtatacgt tttctttttt tcttttaggt gagaaagcag ctcataactc
58921 catagggtta tttaaaatca ccaacgtatc tggtaagaga cgtgatgttg tgctaagac
58981 atttggcaat ggtgaggatg aggatgagga tgacgaagat gacagtgaac gcgatgatga
59041 tgacggagat gaagcttcta aaactccaaa tattcaggta attcttgatt cgttttaact
59101 cttgtgttat tcattcatat ctcttgctc ctaaggcggg agagtctctg tggttgatat
59161 tttctctcaa aattttggta tacagggttc aagggttgct caccatggat ctgttaaccg
59221 tatacgtgca atgccacaaa actctcatat ctgtgtctct tgggcagatt gtggtcatgt
59281 acaggtaaga ttatattttt gtctactttc acagcttggg tgtctttttt tgtatgtgtc
59341 tttgttaatg tggattagct tttgtgtttc ttgacaggtc tgggacatga gctctcatct
59401 taatgcttta gccgaatcag aaacagaggg taaagatgga acttcaccgg ttcttaacca
59461 agcacccttg gtttaacttt ctggtcacaa agatgaaggc tatgctatag actggagtcc
59521 tgcaaccgct ggaagacttc tttccggtat agttatctca gaaatttctg cgatactaaa
59581 taattacatg cttttcggca tggccaatgt aatttttttt tctccactgg ctaggggact
59641 gcaagagtat gattcacctg tgggagccag cttctggttc atgggctgtt gatcctattc
59701 cgttcgctgg acacactgca agtgttgaag atttacaagt aaaactgcca ctttacagct
59761 aaataatatg tttgcttgta ccatttctct ttggctgaac aatctgtttc tttattccag
59821 tggagtccag ccgaagaaaa cgtgtttgcc tcatgttctg tggatgggag tgttgagtc
59881 tgggatattc gacttgggaa gtcccctgca ctatctttca aggcacataa cgcagatgtg
59941 aatgtcatct catggaacag gttctgttcc cactatgctg tataacttaa ttctgtttgg
60001 ttttgtgctt cccacgaatg ttgtagtctt aatttttctt ctttcatgaa ggctggctag
60061 ttgcatgttg gcctcaggaa gtgatgacgg gacattctcc atccgtgatc ttagactgat
60121 caaagtaagt taaaaaccgt agctatatat tttccttagc aatctcttaa caagattctg
60181 atatggtggc atataaatct tgaactaggg tggagatgct gtggtagcac attttgagta
60241 ccataagcat cctattacgt caattgaatg gagcgctcat gaagcttcca cacttgagct
60301 cacttccggt gataaccagc tcacgtaagc aagaatacaa tacacaccga ttctctccag
60361 aaaaaccaa actctgtctt aattgtttt ggctatctgt tgtatttaca ggatatggga
60421 tctatcctta cctcaagact aagaagagga ggcagagttc aatgcacaga ccaaggaact
60481 agtcaacaca cctcaagact tgccctctca gcttctctt ttcaagtttt taactattct
60541 acttcaacta tcattagctt ggtttatttg atctatgtat attttgtctg aaaacctga
60601 ttattttttt ggggaaaacc agggacaaaa agatctgaag gaacttcaact ggcacaccaa
60661 gattccgggg atgatcatct caactgctgg tgatggtttc aacatcttaa tgccttacia
60721 cattcagaac acgcttccgt ctgagctacc agcctgaaag acaaggctctt actctgaaaa
60781 cttctgaagt attactcata gtttgtgtt gtcctctctc tgttctctt ccttgtatca
60841 ttgatggcaa gttgcaagaa ttatacagct ttcacaagt ttcaagtttt catcatcagt
60901 ttatttactg atttcagttg tgacaaaatg tcaaattttg atttacattc tctctccag
60961 taacagaagc cacaagccta ataatagttt aatccttaac aagtaaaaaa aaatcaagta
61021 gattaattgt tgaattctgc tcataacatt ctccaatttg aaaattacaa gtagtggata
61081 aactttgatt aaatgatcga aaacactcat gtagaatgta acaactatca tgacgatgat
61141 gagccatctc ccaaagacaa ctaacaaatt gtgttgttca acaatcgag tcttggtaaa
61201 ctccaccgct atggaaacca attccttctg atgtttggtg taaccatgat tagctccttc
61261 cacaactctc agcttatgat tcggtataac ctttgcaaac tccttagcat ctccccagg
61321 taccacggta tcattcgaac catgaactgt caaaactttg cattgtttat caatttgag
61381 acaagcttga tgcataatcag tgtttaacct atccattaaa gactcttgag taacacgaaa
61441 acaagatttg ccctctgtag catcaataaa cccctgttcc ttaatctttt ctatataacc
61501 atctccgaga cgcacatcat tcttaagatc aaaacgtcca gagatattga cgacattgag
61561 gatataatcg ggaaacttgg aagcatagag aagaacgaca taccacctt tactatgacc
61621 aagaatgaca ggaactagac ggttcatgat gtttagaggag gacaagtgtt gaataacgta
61681 atgcagatca tcctctgctt cactgttgaa gttaccataa tagaaagttc cttcactatc
61741 tccgttacca gagaaatcga aacgaaaaga gctgattttt tctttctcca aagctgtagc
61801 cacgtttttg agaattctgt tggctctgtc cgatctaaag ccatggcata agaccacaa
61861 ttcttttgat ccagtttctg gaagcagacc tcttccaggt gtaatcgtag agaagaaaac
61921 cacgatcttc gttggcttca tcttctttt aaggtaatgg aagtatggaa ctggaactct
61981 aatgaggccc cctattgatg taaaaagaat agatacgaat gagactcgta atcactttt
62041 cgccaaagat caaattgcat ttttctgtgg agatacgaat gactattctt ctcatgtcga tttcagtgg
62101 cttttttttg ctaaaactaaa actcgtaact agtagactcc tcaaccatgg tgaaaactcc acagttttaa
62161 aaatttggac cttactcgta ctagactctc aacagttttt gttttttgtt tgaatgaatt ttacacacaa
62221 agtcacattt ctattgtaaa aacagttttt gttgatacaa caaaaaaat aacacgtgaa gacttatgag
62281 ttagaaaata aaaagtgttt gttgatacaa caaaaaaat aacacgtgaa gacttatgag
62341 ctaacaggct tcccgtttag tttagggaaa ggatcttgct tgcttagaac taccactttg
62401 ctcttgtgag ctaaaataacc tttcacaagg tttttgtata tcaatatcgt cattatacat
62461 tccacctgca aatacagtat tattttgaca actgtgagtt tatagaagat ttcaatttta
62521 gaagaagtca aaagcaatga atgggatggt attgtagaac aaacctcgtc cagatccatg

```



```

54061 tgcagatact aactggggcac caagataatg ctgaatttgc tcttgccatg tgcccaacgg
54121 aaccctttgt gctctccgga ggtttgtgtt tctgtaattt gtagagtcca atcctgtggt
54181 ttgccagttt ctcatacaaa agttcttctc ttaggcaagg acaagtcagt tgttttgtgg
54241 agtatccaag atcacatcac aacgattggg acagattcca aatcatctgg atctatcatc
54301 aacagactg gtgaagggtac tgataagaat gagagtccca ctggtggccc acgagggtga
54361 tatcatggcc atgaagatac agttgaagat gtggcattca gcccgacgag gtaacttctt
54421 agaacagact ccttctattg atatcgtgtt tgtttatgca tactgcagat attttcatga
54481 ttttctaata atacttctgg tgaactttta taccgtgaag tgcacaagaa tctgcagtg
54541 ttggtgatga ttcttgcctt atactatggg atgcgagaac tggcacaac cctgtcacga
54601 aggtactcta tcttttgaat cctatcaaaa gtttgaagat ttacctctt ttgatattat
54661 atcttacttt tttgttttcc aggttgaaaa agcgcatgat gctgatcttc atttgtttga
54721 ttggaatcct catgacgaca atctgatcct gacagggtat ggagaaatag atacaaatag
54781 atgattaata catacttagt atctaattaa gaaattgatg aatatttcag gtcagcagac
54841 aacactgtcc ggttggttga tcttaggaag cttaccgcta atggagtggg ttcgcctatt
54901 tacaaaattg agggacacaa agctgctgtt ctttgtgttc aggtataatc aacttttttt
54961 ttttttttcc ttctttgtat gaagtatact tcttaaccca ctgacatact cttgttattc
55021 aattcagtggt tctcctgata agtcatcctg ctttgggagc tctgcagaag atggtctctt
55081 gaacatctgg gattatgaca ggtgtgttac atagtctact cagatgtcta aaattaatct
55141 ttcttactta tcatcactga aacatattca ttgtactcat gtttgggttg ttttaattaac
55201 catcagggtca gtaagaagtc tgatcgtgca gctaaaagcc ccgctgggct cttcttccag
55261 catgctggtc acaggttctt aaagacttat cttgattttt cttgattgct ttctcatttt
55321 acttgcttct aagttccctt gtttataaac catattaggg acaaagttgt tgatttccac
55381 tggaaatgctt cagacccttg gactattgtc agtgtttctg atgactgtga gactattggt
55441 ggaggtggaa cattgcaggt aaccttgaaa tcttcttggg taccttgata agcaatttta
55501 ttgacatacc gttaaatgtt gtttataact tcttctatgg cagatatggc ggatgagtga
55561 cttgatttac agaccagaag aagaagtcgt ggcagaattg gagaagttca agtcgcatgt
55621 tatgacttgt gcctccaagc ctttaagagta aagaaaacc attgtctatc tatctatcgc
55681 ctatggtaaa ctaatgcggg ttttagcgag gagtcttggg ttttgtaagg ctggtttgtc
55741 ttttgagata ttggtggtag cttttaggac ctttccatat cagttagggg tacatggttc
55801 tggttcatga tctgttttca tcagactctt aggtgctgtt ttgttcaact gagatgttaa
55861 tcaaatecga ccaactttat gtgttttggg ttaaggtttc aataactggc ctgaacctaa
55921 tgattccttc tctgtaacta gtcgagacca acccgctac aaaaaaata gttgcattga
55981 tgttcaaatg caaagcagaa ggggtcatta taaaacaata tagtaataag catcgcaaaa
56041 tttgaaactt gggttcttga tgattcgtcg catcttaaat aatgtttgca ttgtcgtttt
56101 cattagaaaa aaactgttaa tcacgcttag cagtaggatt aggtgatttt caattcacag
56161 atttattcgt aggtatttga agaaatttta agtattaaca gtcactaaac aatcagacca
56221 ttgcaaatca tacatagatc acctggtcat ttcataaaac taaactatat gataatgttt
56281 tactttattg ggcttaacgt taatggtcct cccgcttaat aggttttatg aagtctcttg
56341 atgagctaga gactttaaaa atccaatcac aaagcctaca tgagaccgtc aagatcagtc
56401 agagaagatc aaacgaagaa gttttgttc caaaagatac taaagagaga aacatgaatc
56461 aatgtattgc ttgaggaata aggtagggaa gctaatacat cacaagtaga cgaatcggag
56521 tatttcataa aattaggggt tttagaccta cattttcttt tctttttccg acgatactta
56581 aattgctcga gaaactctat tagtgagaaa tccccatcg gaatctcgat ggaatcatct
56641 ccttctggat ccgaacctcc gcagaaagtc gtttctaaac tgcagaaagt aggtggcga
56701 gctacgatga tcttcaatct cggttttgca ggttagagaa atcttcatta ctttcttagt
56761 aaggttttta gttttacgtc cgtaatgtgt tcatgtaat gcctagctga gaaagtctt
56821 tctttttctt ttatttaatt ttgagaatag aaaatacgaa ttcttctagc ggtcgtctga
56881 tatatatgat actacaatct atcccttgaa ctgttgttgt tgtccttgc tttgattctc
56941 aagaaagctg attgcattac tctgagagt tttgattctc cactctatta tcagatatta
57001 gaacatgctt attgaaagga catttttttg atctgtatct ctgagtcact gttgatgttt
57061 tcagcgacat tgaattaagc aaatgtagta tctagaaact atctatatct ctgtgtgcgt
57121 gtgtgtatat atcttggtca agcatttttg cgttgaatag aggattcaca caggttcttg
57181 taaattttct ggttcatgga catgattgca catcttatta tcttgtaact aatacttttc
57241 aaatgtttca gcatctcgat ttttagctgg cagaactaca agcttggctc aatcagaat
57301 tgagaaaagc atgcttgcct gatttcttat gatatacat tcatgataga caacttcaat
57361 gtttgcgtgat ttgatcattt ggttctgcag cttatatatt tgcgataaaa cgagaaaagg
57421 acattgatgc ggacgagaag aagaaagtta aaaagggcag cgaggctaga cataaagggt
57481 tgaaaaaggg tgctgttaac accgaaatcg agaagaaagg tgcagaagaa actgataagg
57541 ctaaggaagc agaaactgca ataccggaga aggaagaaac caaactgatc cctgaactgg
57601 atccactggt tgaatttaca gatgcactg atcaatccat gtttcaaact gttggcaactg
57661 aacatgtaaa ggtagcaagg aaaccaatc cagaagatga gcaaaaggag cttttcaagt
57721 ggataattga agagaaaagg aagatagaac caaaagacag aaaagaaaag aaacaaatcg
57781 atgaagagaa agctattctg aaacagttca ttcgtgctga gaggattcca aaacttctac
57841 ctgatgattc cgttgattct tcacttctgt attgggacaa attcttctcc aagtagaacg
57901 aatacagaaa ctagtgtgta cttgtttttt gttagtacac caaatgacca ttgttggtct
57961 tttagtttta ctactctgat tgttactcta atcaacgaat agtttaaatg tgatttcttg
58021 ggggttgaga gtggcacttt tataaaactt tggggtaatg tttctatttc ttataatata
58081 aaggacttaa aatataaatt caataaaaat aaggggtgtt ctgagaaaaga gggataaaat
58141 agcttaaaacc ctagagaaac tgaagaatcc taaattgatc atcgtcgtcc tttgagtaat
58201 ttagaaaaatc aaaatgggtc gcaacgtcaa aaccaaggca aagaggaaga acaagggttg
58261 aactttttga atcccaqctt tatcttgttt cgattctcca atctgatgtt ctcaaaccaa

```

```

62581 tccatatcta gccatcttag tgctttggca atcccttcaa gcttcagctg gtgagctctc
62641 gctggatcac tcagcttctg gtttatataa ctgaaacatt gataataata agagtataaa
62701 actcaagaat cttggatggt atcatagctt gtctttgctt acattttctt catgagctctc
62761 tggtagactt ggagctctag cttttccaag acaagataca caccgacct caagaacctg
62821 tataaaagaa gagtattaga tcgagatatt aagttgttcg gtgatagaat caaaatactc
62881 atcatataga gacttgcaca actctactct cgacccaact gcaaatccca acccctagac
62941 aaatccagaa catcccaact ctttctgccc caattttctt ttaacctggt agctctctac
63001 ctaagaacca ggtgttactg acctgtcagc agagctatat aatccactac aagcagccta
63061 tactaattcc aatactcatg ctctagagta tccacaagga ctctactatt tggattagat
63121 gagaaaagca agatatcgaa agaaatacat accgatcttc atgttcttga agagcatgcc
63181 ggagaagcct tagatcaccc tttctcagag cttgcacaat ctttgtgtac tggacacgta
63241 gagaaatagg attaagagag atatgatatc actcagcaga tgcgaaataa gttttgacat
63301 agattcgaag aaaattttat agaacaacag aaaagtctgt acctcatgca gattataatt
63361 tcgcaggagc tcatcttttg gtataattcc taaagaaagc ttcaccggta ccaaatactt
63421 taatatcatc ctgattcaca taaggccaca ttagctgatt aagcaacgca atgcaagagt
63481 aaggtataca cagagtaaaa ccagaaatcag accttatatt tcgttctctt tgggggttgc
63541 aattttgcaa ggcatatgat agctttgtat cagcctgcaa gatccaaaaa aagagattaa
63601 atgttactat caatattttg attgacacta cgaaatgaac tgtgatggta agaaaaaatc
63661 tcttagtgac tcacagcagg aaaattttcg ttgaagactt ctaatcggcc ggtataatac
63721 atgtatgtaa cctggtgttt gcgaagaata aacagaagat aataagtaac cgaataagga
63781 ggccaaaaac taaaactggt tgcagagaac tgcgccagac cttatctctt cttggaaact
63841 cctcaaagtc aaatatccga gcagtttcga tacttcttat tacacttcga caaagattga
63901 cggttccaag ctaaaaatac aaattgggga tacttgttta gtactccatt taatgggttaa
63961 atcataata tcaaactggt gccagaagag atatcacctt aaagtaggtc ttgaacattt
64021 ggcaagtcac atataatgct ccgactcggt ttggaccttt cccctatagt ttaaaactac
64081 ataaggagcc ggaaaaaata ttgaacacaa atataaacta atattcctga agatatatag
64141 atagatcaaa gggaaaactt acagcaagaa ctccaaaaac tttcatgaga agagaccag
64201 ctgcttttaa tttctctgga gactttccat tagaagttaa atctttatca gcctgcaatt
64261 ccatgtggaa aaatcatgct ttcagaactg actcctcgac acaatgaaac cagacaaggt
64321 taagattttg aaaaagggat atacaaatga gatactacct tttccgcaag aactcgaatt
64381 tcatacaaa ctacatacag agcttccaat gcccaagcag attcccaatt acgaaactcc
64441 tgaacaaatg cactgcacat atagaatcca tcagatatca gacatttga acagacaaaa
64501 aaaaaagctt atcacctga aatcatatta gacatcaata gaggagcaga gaggaacaag
64561 aagactttac ttggcaaact tatcgaaagc aaggtatgct tcgactagat tcccaactcg
64621 gtaactttgt agagaacgaa acacatgagc cagaatctcg ccatattcag agaactctatc
64681 agactgtcta atcaaaactgc tagaatcctg tatcagtcaa agaaaagaaa actcaaagga
64741 ttcagcactt tcacagctct ctacatataa ataaaaacc taaatctaaa atccatactc
64801 ataaactcaa acatagacat agagtgttca ctacaacaac tcaataggag acactttcac
64861 ggcctaaatt cactactcatt ggtgaagcag cctggaagac gttgagtgcg tcggcgagtg
64921 gtgaagtagg aggatcctag attgacattg attgaaccta aaagatacac actttgctaa
64981 cataaagata atgttcttag tgttttgagc cctaaattca gaaattcata ctcatgtgtg
65041 aagtaggagg acccaattag aaaatcatat tcttaaggat cttacattga atcgatttaa
65101 cctaaaagtt acaaaactttg ctaacataaa aaccatgctc ctagtgttca caattcatatc
65161 gagaacccaa tcggaacata tcagaagcat cctggatttg aaatcgaacc taactttacc
65221 aatttacaag gatcatgcc aatatcacag gaacatactc caagctgcta atttgaatca
65281 tgggtctcga ttaacagag aatatatgta cctggaagac gttgagtgcg tcggcgagtg
65341 agagaagagg tggagaattg gaagagaag agaggagacg gcaaagtgtg gaagagtctt
65401 ggtacgaaac ggcgtcgcag aatcggttta ggtattcagt gattctccgg tgagcttcac
65461 ccatactaac gtacgccatt gtcaccttct tcgtctgcgt cttccttatt ctttgagaga
65521 gacacgaaat caaataagtc gccaggaaaa aaaaggaaat aacggagtgt agaaaaacga
65581 cggcgttttg taattaacta aattgttact acgtactagt cttttctggt attagactta
65641 ttgacatatg ttagttgggc catttaaggc ccatatgaaa tacaatcatt aaaatgttt
65701 aaaatttatt tgtaaaggat ttgaaaatat cattagtgtt ggaaggattt taagtctctt
65761 gtaatcggtg gtactgatca tgaatttttt ttagagtttt taagttcata taactgtttt
65821 atagatcatg attaatgttc ttttaaggtt ttttaagttc tgtaatcgtc cgtaatgttg
65881 acatctgtat gtaaatatgt cttcttagcc cgatggtgac gtgaggaact ttcttgagg
65941 gacatgaaat caatttatat catgcatata actcaacaag ttgccaaaaa ctcgagaaa
66001 tgaatataat caacttgtca aaagtaatta actgatcaaa taaatcactt gtttccaaat
66061 gtgatttata ctttcacact tctatatatt cccattgata tccccctctt ttcatagtta
66121 gcaaaggcct ctactacatg tacatgataa ataaataaat aaaaaacacg ctaatcatat
66181 aattatcaat atattttatt ctggaacaca caattaatag aaattttgag ttgagagaag
66241 tcaatccagg attagtattt atgtgtaatg ctaaaagttg aagataacga gaaacttgag
66301 aggtttttta ggttcctact ttatttttgg tatgaagggt cctactaaaa tatgctctta
66361 tatatcattt tatttaaaaa gcattccatt tgaaactaat ctaaaaatct ttagattcat
66421 tgttcaccac tcatagtaag agttgcatgt ggaccatctt atattctcta ctttctcata
66481 cttcgtattt ggaaccttta ttccaacaaa ttcaatgtgt aaagccatca ttacgtttcat
66541 ttatttttgc ctttttgcgt acataaaaga aacatcataa cttgacatgt aaaaaaaaaa
66601 atcattgatt cgatttgaga cttttgaaaa ccttgtcaat tttataaaat ttggccgatt
66661 tgatctaata cacataatcc atcccttata gacgaaatta aattttcaaa ggggatatta
66721 aaaaaaaata tcaaaaaatg gaaaaattta atatcaacaa tttttttttt aacaatactc
66781 cataaacata ataattgtat taaatttaat tttttaatac aataaacatt ttttaacaaat

```



```

66841 ttaaaaatat aaacttttct tattctctat gtgttatcca tcacatcaat ttcaacactt
66901 tttttttttt cgtcaatttc aacactttta agaattaaaa agtctcattt tctttttatt
66961 cctttttttt taaaattcct ttatttttcg gataattaaa tgattctcta agcttcatag
67021 caatgaaaca ctttgaattc gcatttcggc gatatagtct ctaataacaa cctcgccgtg
67081 tgctcttgcc tcaccaccgc atcttctttc tccaccagaa ccgctccgac gatcctctcg
67141 tatctctctc caccaccatt agccacataa tcaaccaacg ccgctgtgac tgggtccata
67201 ctaggattat acgccgcgca ttccataata caacctctgt acactttctc gtcacaattc
67261 accagcgaat ctcccgatgg acataaaacta tacggcgcggt acgatctatt cgccgcgct
67321 aaagccgttt gtttcaaatt ggcggatgaa tcggtgtttc cgttacaaat cgaatccaga
67381 tctgagattt tgagatggtt atcgtgagat tcgagaagaa gaggatgatc tttcccgaga
67441 agatcgtcgg gaccgaatct gtgtggcaag aagcttccga gacgtaagaa tccgtctgaa
67501 tcggcgccgg aatcggaatc ggcggagttg tttggatcgg tgataaggat tttgatttca
67561 ggtgcgtcgc gaatttcttg gaggaattga cggcaatggc cacatggtgc ggcggagacg
67621 gcgaagaaat tgagatgacg ttcaccgttg agtgtgagat tgggtgacgag gaactgttcg
67681 cgtgtggttg agtgggtggg agggagattt ggggaattcg cattgacgcc taagaagatc
67741 cgacctgatg atccgagtcg gacgactgcg acgttgaatt tcgaaatcgg agttcgagcg
67801 taggattgtg ctggtttgac tagcgacggg aggagctgaa tgacggaaac gccgagttgt
67861 ttcgcgccgg attctgcttc tttggattgg attacgaagc ttggcttatc cattgcccgt
67921 tgatccggtt tgggtcgcgt aaatgggtcg gggtattttt agatgggaga gatccgggaa
67981 atgaggaatt atgaaatgga tgctggacaa acaaactatt tatagataca acaacgtagg
68041 aaactacatt acgcaaatga agttttgggt cggtttcggt tatgtttcgg tttactgttt
68101 gtacaaaata gaaaaattat atcagaattg aatttacaaa taatatgatt tcggtttgac
68161 taagccgcat ggaatcattg ttatgaaaac taccaaaaat gcaatagagt attagtgtaa
68221 cgtaatgtta aaatcgtggt ctaattaaga aaacacctag tagtttcatt aaccacgtta
68281 cgaaactaca ataactctcc gataatatta ttttagaggt cagaaatgaa ttaaactata
68341 gctaaaagac agcttagttt catataattt tgatagtatt aaacatttta attccggtgt
68401 tgtcaatttt aagttcccaa aaaaaatggt aagaacaaag gtgatagtac tattgttgac
68461 aaaagaaaaa aaagattagg tttgttctt atcgttggaa tcaacgtcac agaccaaaag
68521 agaagtcctt gttgataatt acgttttaac caacggttaa gtttttctta atagaagaaa
68581 ctgagaaaaa tctatgacaa caatcagttt tccaattagg gtcagttggt tgtttacata
68641 gataagtgcc attttgtcga acataaactg ataaacatag aaacaaaatc cataaaaaaga
68701 aaaaaacctt ttccacacac aaagaaatca taatatgacg agctatttcc gaaaactcca
68761 atctcaacaa actacaaaca aatttataga cacaacaca ttttagacc ccaataacata
68821 gtatttgccg aagagatcct agtctgagc attcctaaac gcgctacaag caataacata
68881 aaccatattg agaacaacga ctgtgatgat aagtatgaga tttgctttac gccattcttt
68941 tctaaggttt cccaataaac cagctttgca tgagttgcaa ttgtaacaaa gctggctttg
69001 gtcattgctc cataagtaac agtctgcgtc tgacgccata ttggttggat ttggccacag
69061 tgttgggttc acaaagttgt agccacatgc ggttgggtgt ttgcagcagc cggactgtaa
69121 aacaacaaat gttcaagaaa ttgtcataat tagatttgac caaatatatg ctttaacctg
69181 ttttcaaata tgataatcat agtcttaaag atattttacac ttttttttta ttggtaaaaa
69241 atattttcac attgtaaaca caaacgctac catttgcgtc tgccctttgt gtttacaaca
69301 aaaatttcat tttgcttaac atggttaata ttatgggttc atgtggagaa aaaaaaagta
69361 tatactaaag gcaaacacaa acaatcaaga gttcaagagt caagacctaa aacaaatgaa
69421 cactagtttt agtctgaaag ctataaaaaat aaaacgacta caaagaaacct aaaaaattgc
69481 ttgcttggta tgaaagtaca taaaaaaaat gaaaaaacat gttgaatctt atcacagaat
69541 aatcttcttt tactttatgag taaattttgg ttctgtctcg aggacgagaa attgtataac
69601 tctttcttag aaataaactt atttttctt gagtcctaatt tatccaata agaaaaatta
69661 taatgtgagt ggggcaacat gggagggata taattatcta ttatgtttca tgatggagac
69721 aaagacttgt gatcgtaatt attgaatcaa atcattggtc caatatttaa taatgtctaa
69781 tgactaaaag attttaaatt cagtcgagac ctgagcatct aggccttatt aattattagt
69841 ggtttccaat ttgttacagt gacaattggc taggatgaca acttgctata gaagttttag
69901 gaataaaaatt attttaaata ccttaattct ctccatgttt aaagcttaaa agacataaaa
69961 gacatgacaa tgaattttta gagtgatttt aattaataaa attaaataga aaaagtagga
70021 aattttgtca agtttgagtt gtgatcacga aaaacactat ctaaagttaa tgggatcatt
70081 gaattttatg agtgagatag tgacaaagga aaaatataat taatacacag ataaaacata
70141 gttttatcatg taccatattg tgggtgcatg ggtgggtcct tatgatagtg tataattttg
70201 aaatctgact atcttactat ataactttgt ggacttggtc tccagcaact tctacaccag
70261 tctactagga ataaatatgt gttcaataat ggtgattaca tatgcttaat ctatgagttt
70321 taccacaaaa tggcactaac tgcgaaacgt taagaaagct aatgtagggt tctctccaac
70381 actagaaaaa ataagattaa ttaagtcaga aaatatTTTT ttagtttttt caaagatgag
70441 gaaatccaaa acgaaataac ttaaagctgc aaaacaaaaa aaaagaatat agagaataag
70501 aacaattctc ttttggacat cttcacaaaa gctaaagaga aattttttta ctttttgatt
70561 taaaaccaca tacttttact tttgaaagaa aaaataacct atctttctat ggcaagaaca
70621 aaccgtaggt ctcacataat ctgagatttt tacaaaacat tctctttact cgtatcttta
70681 tctccattaa tcatctataa atactataat aaattcgttg tgattcgttg gcgatgttgc
70741 ataaccgata tttagtctag atttacctct aaaccaaac taaaccggaa actagaaatt
70801 tcgggggttct ttttataaaa gagaataaaa gaagaaagaa catgcctgga gaggagtgat
70861 cttagaggag gagaagaact gatcggcggt gatgaattct tggttgagtt taggacaaac
70921 attagatca gccaaacaag cccttagcct tccccagttc ttggaatcca caacgttctc
70981 cttaaagcaa ttcgagaacc cttcaagcct atactcttta taacctctac ccggaacccg
71041 atacgatcca tcgggcccgg tcacgacgaa tgcaaatata agaaccacca gcaaaagtcc

```

```

71101 gatcaatata gccatacagc acaagtaaac cgccagtaga gtttccttgt acctgaaata
71161 tgagagctca aagtttacat ctttatgatt caatttgtgt ctatgttttc aaagtcttga
71221 agttctttga ttcttacttg taggcgccga tgaagcctgt ggcggagacg acgaggatga
71281 gaacgccgag gacgacaacg ggccaacgga ggagattgac acactcgttg tctggctttg
71341 aagctagcca tatacctgac gccgttattg gtatggaaca gagtaacgct agtaagttga
71401 gtatcgccgt taagttattc gctaacgcca ttaaattttc tctctctctc tctctttgtt
71461 cttgaagatt cgtgcttttt atggaaagaa aatgaaataa tgagagttag aggagtgtct
71521 ctgtttttgt cttaatgact tatgaggtcg gtggttagag atgatgggat gtgtgaaacc
71581 agttactgct atttacagaa aacgatgttc ggtttggtcc ggtccggtta atcactgatt
71641 ttgctgattt ggatttttctg acaacggttg ttttaagaaa aaagaattag atattggatt
71701 aaagttaaca tgtgaatgtc agaataaaat atgttttagc atgaaaagtt aggttttgaa
71761 caagtttttg gttgtgatgt aataaagaga acttgtaaat ctggtcggac cggcttttgg
71821 tttagaattt ggaatcttga acgggacacg tggatacgcg agatgccttt tgtgttaata
71881 ttcttccatt ctgagagtta ccgaatgcat caaccagtga gaactatcca cgtggcaccg
71941 acgacttttg gtcatttgca ttgaacgtat gttactttcg acagagaatg caaaatggag
72001 taagttccat tgttcggaaa tcggaatttt cttcttttcc aagcattaga ggtttcggag
72061 caggatttgt cgttggtcga tgagtgaatg gtcgggtaat gtcgggtaag taagcttttg
72121 ttttggttca gaggattggg attgcacaca agctgctcga tagaatgtgt gaacccaaac
72181 atatggatga atttttctca tctcgatttt attgcgattg gagaagaatt ttgagtcaat
72241 acacacttag ctttcagttt tagatgctga tatagaaatg ctatttgttt gattctagga
72301 attgggaagg aaatgagtca aatcagagag aatatggaaa gagatgatgg cggagtgtgtg
72361 gattacgtta agagttgaat tacatcttat aaatcatgaa attgactgtc aatcttagag
72421 cttagattac aagctctttg tctaattcga agttcgacac agagcttgat gagactgctt
72481 cagtttggct ttggaacttt tcaagtgga tgaagaaag aagagtgaag tagtatagct
72541 gaatcagcaa gattctgaca tggctggtta tttaaattat aaggaaacac aagaagaaga
72601 caaagttttt atgatgtgag tacagaagcg tcaaattcca gatgcacctt aggtgaatc
72661 cgcatttccc attttcttca tggctcgaag tctggggcct ttgtctccaa acttggtagt
72721 tgaataaagc tttaggtagt cttgaaaacg gtagccactt ggatacaaga gctttggagc
72781 tggagatatt atggcatcac cagccggatt gtaaaacgta gcaatcgaca gtctacttcc
72841 atgcttcaat gtcattacac ggtgaacaac actcttgtac ctcccattac tcagtatctc
72901 tagttgatca ccggtattga caaaaatggt attgttcttg gatggcggta taggaacca
72961 tctcccatct ttaaagaact caagaccagg cacttgatca tctgcaggga taataatgat
73021 tccccagca tccgtatgtt ctctcagccc tctcataagc tcaggacgtg ggcatctctg
73081 gtatttagcc acttttgttc caaaagctgg accttttgg ccagaaaagg cattcattat
73141 gtcttctga tcaagaccaa gattctcaca catgagcttg gagagtctct ctgcaaactt
73201 gtgcagttga caaacatatt catccatcgt cttgctgcaa atatcaattt caccctctca
73261 cattgttttag ttaaataagc aaacaggtaa aaatggtaga gcctttggta ctattcaatt
73321 cttaattact ttacctgagt tctctgaaa tgtttgggat ctgacagata tttgaagttg
73381 ttttatgtga gatgaagaaa ctgctttccc aatctgcata tgaggttttg ccttcaactc
73441 aagccttgac catctctgac tggtaaaact tctctttcaa atgtctctca tagtgagagt
73501 taatcatctt cttcactttc tccatcaact ctttatcaat tccatgatta tcaacctgaa
73561 aattcactaa ctctttaagg atatacagta gacatatctt agaaaacact gccttgtgca
73621 taaagatgtt atacatacca tgaagaatcc ccacttatca catgcatgat caagaagtga
73681 catggtcttg cttctctctt ctccatccaa ctctgcaaaa tcaataactg gaatctccat
73741 ctctctctct ttgatcaaaa ccactctctt tttatttact ttttctcaca cacagatttg
73801 ttgaatgcaa tctagaagaa gattttatag gcataagtga aaattaaatc ctgtgttctc
73861 ttgtgtgtta ggaggcatag ttgcctactt ggaggcaaaa acaatagttt ctttcaacac
73921 tattctgtta actcattgaa tttgaagcat cattacgttt aatgaataaa aaattgtata
73981 agaactctgt ttggaaatgt aaaatgaaca ataggattga agaaaagtc aaagttagtc
74041 catacaacca tgtgatcaca agtttttgaa ttttccataa tttcttaatg gttttgggtc
74101 actcccattg tgtttccttg taacatcatt tatgtgaagt ttctttcttc ttcacaagga
74161 aagttttttt tttttttttt tttttttctt atccctactc atattcgatg gacctagcca
74221 tcatgtctaa ttgtgtatta tgtttttctt ttcactgatt ttttttttct ttaaaacttc
74281 actaaactac ccatctaagc gtcattggcg tacttgtaaa atgtgttatt ctggtgtcac
74341 caaatttggg tatgctgtgt gttgtgttta gcaaggagct ctataacact catcaaatga
74401 taatatgacc cttttttggg atgaagaaat ctacttgcc taaagctaca gtatcgaatt
74461 gttacttgag aatagatttg tgaagaggta cagtatcgaa ttggttattt tgagattacc
74521 ggagtactac acaccatttg caatagaaaa agaaataatg tgggtctaag cggaattaa
74581 cgaagaatg ggccccctct atgggcgctc tcggttttga atatataatt ttgtgtgtca
74641 tattttataa cactttgtta agaattttgg attgtgttta aagaatgta tcattttact
74701 ttttgggtgc gaatatctcg gatattcgtta gttagttag ctctagtcta tcaaaatcag
74761 caccgacaaa tatatttttt atgtgtgtga aaattcaacc atgacggccc cttagagcgt
74821 aaacgggcca agacctaatc tacaagact tttcagttat atataacttt gtttcgttta
74881 gtttgactgt ttgaggaaga aatggtaatg ttattgaatc ttttgtttgg gtcatttggg
74941 gttgtaatat gttacataat tcgaaatgat ggtttatcgt taacacggac tttaaatatc
75001 attaacatgc gtgtatttga gattacaacg aattgattca acatttcaac ttttaacatg
75061 ttggagttaa ttattcaaaa agttataacc ttgtcaagat atctaaacgt tttcaaattt
75121 catgctattg atattagtta ggttcataag acctcttag agtcttcata attgttttta
75181 gtaaccccat taaaccttct gcttcaccaa gagcctttct ttccaattta ggtacatttt
75241 cttttaagta tgagaatttg gtaaaactta actccaaact aaacaagcca aaagagacca
75301 aacattacaa tgaaccttaa acaaaaatta taaaatgtt aaagtgtaat acgcaatttg

```

75361	caaattcagc	aactatcacc	tttactttgt	tcacttttgtt	gtataaatga	tttttgttct
75421	cactccctga	accacgcgct	aacggtggcg	cgtgttctta	acacactttt	tagtttccag
75481	aatcgatgga	gactttaatt	aacattttgc	ccttaagcaa	acaataaaca	ataataaaaa
75541	aataaaatta	atttccctct	ctctctcaat	ctttgtgctt	tctcgttctc	caaccaccgg
75601	aaagagagag	agagagagat	ctttgtgtgc	ttcttctact	tcttcttctt	ctctctatct
75661	ctccgcccgc	ccacattcac	tgagatgcgg	cgggtgaaaa	acaacactga	caaatctctc
75721	gtgataacga	tgaggcttct	aacgcttctt	ctgatctgta	ctttctctgt	cttctctctc
75781	tttgcttatt	ccgccgagtc	agataacgag	actgactcag	ttgttacacg	tgagatcaat
75841	ggaaccgtcg	ttgagtccaa	tgccacgagt	gcgaaaccta	gggaagatag	tttcgctgat
75901	atgatcgatc	gagcacttga	gaaagagttt	cctgataatg	accagaacga	aggttcctcc
75961	tctctcacta	gatagatctt	catttcttct	atttgttacc	tgagctaggg	tttctctagg
76021	gatttcaatt	ttgttattat	tgatccttct	tctctagggg	tttatcattg	attagctgag
76081	tggttttagc	tctctctttg	tgccacatt	atagcaatta	atgcttttgc	gttgtttcga
76141	ttttgcatgt	cctgatccag	gaagcttcaa	taatagtgtt	gctgactcgc	aggttggttg
76201	ttgcaagttt	agatcttttt	ttactactct	tatgttttct	ttctttttaa	ttaaaggcat
76261	gtgttgtgtc	cattaggcgg	ttctagagac	tgttgcacga	gttaagccaa	agaaaaatga
76321	aaccaagacc	aaggaggaga	agtgagtttt	tctgatgctt	attttgcctt	gtttggctgt
76381	ttttttactt	ttgtttcact	ggttaaagca	aactataaca	tatgagtgat	tttgttctag
76441	atccttcttt	aatttggata	acgagaatgg	cgtagaggat	actccaagac	tgatagatag
76501	gaaggtgaag	ttcgatcgtc	ctaaattatt	tgtactatat	attgctgaag	aatgtataat
76561	ctgttatgta	cccatatgta	aaggcattct	cgtgatgtga	atttgcctgc	attttgatcc
76621	aggacaacgt	ttttataatg	tccaatccaa	aatccaagta	ccctgtactg	cagctagatt
76681	taaggtgagt	atagcattcg	taagagaagt	tcttcgttcg	attatatatt	ggaagtatct
76741	aacttacttt	tttttggttg	gggcatctct	cttaaactcg	tacatttgta	ggctgatatc
76801	agatttggtc	gtcgtcattg	tttctgtctc	ttgtgggtga	attgcctttg	cttgtgcggg
76861	tcaaccggtt	agcacacact	tcctttgcaa	agcttgtgac	ttcaatatta	ttttattagc
76921	attgagtatc	cagcttttccg	gcattgtaga	aatgattttg	ctagctctta	ctaatcaaac
76981	catgacagcg	ttatttaatt	tctacactac	ttatggcagg	tcattactgg	gtacttattg
77041	gctggatcta	tcattggacc	gggcgggtta	agctttgtta	gtgaaatggt	gcaggtagct
77101	tacatcagct	ttagtacttc	aaagttcatc	tttttttttt	tcttcatata	cgtttgtatg
77161	ttttcatgac	taccagattg	cctgggacaa	gggcttgatt	tataattttt	ttgtgtagca
77221	caggtcgaaa	cagtagctca	gtttgggtgt	atctttctcc	tttttgcttt	aggattagaa
77281	ttttctgcag	cgaaggtttg	catgttttca	ctctcttata	tttgtctggt	ttctttgtca
77341	atgcgattga	aacttttgag	tcagttattg	tactgcgtgg	ttgcagcttc	gtgtgggttcg
77401	tgcagttagc	attccaggag	gtcttcttca	gataattttg	ttcatgtgct	tgagtggaat
77461	aacagcctcg	gtatgttcta	aaagtgtagc	aatcgagtta	agaatgagtt	attaggcatt
77521	tgtggtatga	catgtatgta	gagtgcatta	gggtgactta	gctgtttctt	cattgcagtt
77581	atgtggcggt	aaactaacag	aaggaatatt	cgtaggcgca	tttctatcga	tgtcatcaac
77641	agcagtggta	tattctaatt	tcctatgtgt	acttcaatta	atattttgat	cttgttctac
77701	actgtttgaa	tatttcccta	ctctcactta	ttttctcttt	tctttactgt	gccagttata
77761	ttttcctaaa	gtattcttta	gacttgattg	tcatatgtta	tccttacctc	cccatgatcc
77821	ttgtaccagg	ttttgaaatt	tttaatggaa	agaaaatagca	taagtgtctc	acatggccag
77881	ataactgtag	gaactcttat	tcttcaggta	tctctacctg	ataactctca	aaactatagc
77941	ttctctatta	tcattttcac	acgttttttt	gtttttccta	ttattctaaa	actgcataatc
78001	cttttgatcc	gaaatatgat	tcaaatacaa	ttccaatagg	tagtatgttg	catgcagcaa
78061	gtacttttgt	tagaacttta	atttactggt	tcatgaatca	ccaggattgt	gctgtgggct
78121	tgctgtttgc	tctcctgcc	gttcttgggt	gcacatctgg	tgtccttcaa	ggagtgttgt
78181	ccatggcgaa	atcgatgta	tttttctctt	ccacgaattt	ctatgaaact	atcgacaata
78241	ccctgcttcc	attgttcact	tgttacaata	aaacctagtt	tttggttaca	ctttcttttt
78301	gactgggttt	taatatcttt	ccaggttggc	tattttgatt	gcgttttttg	gagcttttgtt
78361	tgtattatcc	cgtacctggg	taccttgggt	tctaaaactt	atgacaagcc	ttctctctca
78421	ggtatagaca	catttctttac	ctgctccagt	tttggctctt	aaattgttta	tcaactcaag
78481	cgtttttcaa	tctttgacag	actaatgagc	tctatcagtt	ggccgctgta	gcattttggt
78541	tacttgtcgc	ttgggttagt	ctgatccctt	tgttcttctt	agctattggc	tgtttgcaca
78601	tttgtgactc	tcgaattttc	atttctttaa	tccagtatcg	ttttggcttc	ttcagtgtag
78661	tgacaagctc	ggtctaagtc	tgaggttggg	ttcctttgcg	gcaggagtga	tgatctcaac
78721	aactgatctc	gctcagcata	ctcttgaaca	ggcaagaagt	caagtttgac	tgttttcgaa
78781	tagagggtct	ctaccatata	attttggcca	ctctcttatt	ttgaaagtgt	ctttcttgct
78841	tataactat	cttggtttcc	tgatgaaata	cgtttctctt	ctttggcaca	ttctacccaa
78901	gtttttcata	taaacctgtt	ctgatcgctt	ctgtttctct	gcagggtgaa	cccatccgca
78961	atttttttgc	agcactgttc	cttgctagta	tcggcatggt	gatacatatg	cacttcttgt
79021	ggaaccatgt	tgacattctg	ctagcagctg	tgttactggt	gatagtgata	aagacggtgg
79081	tagttgctat	cgttgtttaa	gtctttggat	acaataacaa	aactgcagta	cttgtaagaa
79141	gcgaaccctt	ttttttttcc	ttctttctca	tcattcgaaa	gaagccttaa	ggtttctctg
79201	cttgagggtt	ggtatgtccc	ttgcacagat	tggggaattt	gctttttgtc	tgcttaagtctg
79261	agcatcta	cttcacctaa	ttgaggtta	ctctctatga	ttattgatgc	ttagtatat
79321	tatatgttag	catatatctt	tgaaccctg	tgtgaccata	atagatttct	cataatttgc
79381	ttttgggtga	gcgataagta	acaatgaagg	ggttttgtgt	ttatttgcag	agcaaattgt
79441	acctcctgct	tctgggaaca	actgctttaa	gcctgggtatg	gtctatttcc	tcatcatatt
79501	tataataagt	tcataattcaa	aaacgaataa	acgaatattg	acacggacgt	tgctttatca
79561	atgaataaca	gtagaattta	catactccag	tatcaatatt	gtataattac	taaacaaatt

83881	ttggaacaaa	ctggagcagc	tccaaactgc	aaaaactctt	gtggaaacga	tggaacccaa
83941	cacaacgtga	gcaaacagca	agggagatac	aaagaaaata	atgaattggt	tcaagacaat
84001	caattctgca	aacctcatgt	cctggaggac	taatatggta	gatgcagaaa	ttatggagca
84061	gtaaggacac	tgcttcaggg	catagaaaca	atccttggaa	acaagacata	tctagcagcc
84121	ataaaaacat	acaagaatct	aaaatcagat	taccaaata	tgttaaaatg	aaaacaaatc
84181	atgtaccaaa	tgtatattcg	ccactcttac	agtttagtgc	tacatctgga	taagtgatca
84241	atgctgggtt	ctcttgatct	ccatatacta	caactgaaac	tgaaccaatg	caagttttga
84301	catgatgctc	ctgcaagaga	taagaaaata	agcaagaaac	gaatcagctc	acatccaaat
84361	taagagaaaa	atataaacaa	taccacatgc	ttcattctca	tatcctcacc	attagtgttc
84421	atccaaaatc	aactcaacat	ttttaatttg	tggtctttta	atcggatata	gctaagaaaa
84481	ggaacattcc	aactttgttc	cagacaacaa	aaaatattaa	ggacctcaag	agattctccc
84541	atgtgcaatg	actcagaaaa	atcagtagca	agagataact	tatggctaaa	aaaagtcaga
84601	cattatatat	caatcatgga	gccataaaaa	gataccattt	tcaagggaat	ctcttgacc
84661	ttttatgcaa	gaagaacaaa	gaatcaaact	aaaagaaaca	cccaacttgt	aaattttcaga
84721	tgacaaagtg	aaacaacttt	tcggtatcaa	tgaagtagta	ttaactctcc	aactagtaag
84781	taaattaagc	atacaaaact	aaatcaagaa	cacaatagac	aaaattgatt	tcaatcacga
84841	aactgaagct	agatttttaga	aatcagggag	ctaattagat	gctatttcgg	gaaagtaata
84901	accagaaagg	aagaaaacga	aatgtgactg	ttcttagtgc	ctaagagtga	agacaaaagg
84961	agaaaaggaa	gaaaatttgg	tagctcaaag	cgtaaacaaa	ccttgccacc	attgcagatc
85021	tcttcaatgt	cgagtgcagc	ggcgttggtt	aaaccaccca	tgacccgaga	aacccccaaa
85081	tagtgcgcct	ttcttgctct	ttctattctg	aaactgaaat	tgaagagatg	agcaacaata
85141	ataaagtctt	aagcttttatt	gtgggtactg	ccttagattg	tacgcttcaa	cggctctttt
85201	gggtttcact	ttgtcttctt	cttcgatgga	atggaccact	actacttata	tctctctctc
85261	tctctgtctc	gatgtctgag	tcttataagc	aaagccgcgt	tgctgttaat	tgtccagcgg
85321	ttaatcaaat	cattcccacc	gacgcgatg	agtcttcttc	aatttgattt	gcaccgatct
85381	ttaaacgycg	ccgtttctgt	aatcattttc	tcatttttatt	catactataa	ttcattttat
85441	tttatctcta	ctaatttggg	agctgaagtt	ttgattgggt	acttgccgat	cttcgtagtt
85501	tttataagtt	acagtataac	ttctataaat	taataactta	taaattataa	atttttataa
85561	attaataatt	gggaccaaac	caatataaaa	attaacacaa	atcgataaat	aataagataa
85621	taattttttt	gaaagttcta	tgtaaatata	tggttccatc	aatatcatta	attaataatt
85681	acataaatgt	atcaattata	tatatatttt	atatgtaaga	aatttctttg	aaatatattt
85741	ctaataattt	tttccataaa	tttgatttta	tttttattta	aaatttagtt	ttaatatatt
85801	actgtatcaa	aaacatttga	tggtgttttc	taaatatgat	atgattttta	cttattattg
85861	attttagaac	aatattttct	ataatgaaaa	aaaatttata	gaaatttttt	aaaagttatt
85921	aaattaggaa	aatctctcta	taaaataatg	aatattaatt	tatcgataaa	ttaaaacctc
85981	tctaaattaa	taaattttgc	agtccecaat	ttattaatat	atagagattt	tacggtatat
86041	tatatcggtt	tgtcaatttt	atggctaaga	agaaacaaat	tgtattgttt	ccaagttaaa
86101	tttatgttag	attaataact	tttaccattt	ttatgctttt	tgacgatttt	tatatattat
86161	cattaatagt	taatcaagaa	aatttttgata	tcataactaa	aaaaaataga	acatactggt
86221	ttttattagg	atcaatcaat	cacattcttt	ctttcatata	attctcttga	acttgaaatc
86281	tttaacgatt	taatagtttc	tcaatatatt	ttttcaaaca	aaacaaaaat	ctcataaaaa
86341	aactcagtc	aacgaggatg	atgaatcatg	atccaaactg	agttttctta	tgagattcat
86401	catctcttag	aaccacattt	ttataactca	taagaaactc	agtcgcagca	ggtgatatta
86461	ttttgggact	cagttttctt	tttttgagtt	tctgaaaaca	aaaaatgcat	aaaagtgca
86521	atccagcatt	ataaatattg	ttctaagatc	agagaggaat	agatgagaac	caagatccaa
86581	actagttcta	ttacaaagag	ctaggtgtca	acttaaaaag	tcctaattaa	agaatttgca
86641	ttagcttcac	atcctctaca	tgatccagaa	aggtgtaaac	attactctct	ttaaatgctt
86701	ccattccttg	catttcaacc	tttttgattg	tggttctttc	gatattgtag	tagaaaacgt
86761	agaaacgcat	acctttttat	gaccagaaga	caaattcctt	tgtctgagtc	attccaacac
86821	aacgtaaaat	tgtgggtttt	attatatcat	gcattgcagg	caatatgtat	gtatgctcag
86881	accattcttg	ccttttggca	tcttcaagaa	cccacaattt	aatacttctt	attactccat
86941	taaaacggga	ataaactggt	gacggcagat	acaaacttag	tttcccattg	tagtttatga
87001	gagatccatg	atgcagtgct	ccgataataa	atctaaactc	ctcagacctg	acgtcaaagc
87061	aaactatcac	aatatcattg	ttgttgacct	cagcataata	atatataaga	ccatcaatgc
87121	atatccctcc	atcgaaagaa	cgatgtggta	tgcaacattg	gatcgttctc	catgacatgt
87181	ttccacttcc	taaggtcaaa	atatgatgca	tatgataatt	gacgactaca	ggagacgact
87241	tcttatgac	atcatcagga	ttttgaagct	gaggtgacga	taaaaataaa	taactcgtac
87301	cttttctgta	ggcgaacaat	agcttcgggc	gagccaaaaga	tctagtcaag	aacaactcgg
87361	tgaaatatgg	acgggctaagt	gtggaggccc	attgcttcga	tacgcaacga	catctcgcta
87421	tagaattcac	cgacaacctc	aagaatatct	cgataatgag	atcaattggg	atctgcaagg
87481	agtttttagt	tccatcctct	aagegcagag	acgttttcat	ggccgaaata	aataaacctt
87541	aagagagaga	gactccttgt	agaagaagaa	accaatgcgg	taaaacaaaag	gcaagtaaaa
87601	gccaaaaaaa	agttaatcac	aatggtaaaa	acatggactg	agagaaaagg	cgtatagccc
87661	aaagtttcca	agcttggtata	aggcccaact	aaataaccta	acgagtttaa	agttcaaacg
87721	tgatgaaaac	ttaccgtttt	agcgtttctc	atgtttcttc	catataaata	gttttagttt
87781	tgtagaaaac	cctaactcgac	gacggccatt	atgataaatg	acggcggagc	taaacgggag
87841	acgctgctta	gggttgacga	aatcggtgga	agaggaagga	gtttgggtgg	ggcacagtct
87901	cttcgtgctg	gacaagttat	cctcagagag	tctcctctcc	ttctctactc	tgcttttcca
87961	tttctctcct	cctctgtttc	tccttactgc	gaccattggt	tccgtttggt	agcttcatcg
88021	gcgcatacga	aatgtcaatc	ttgctctctc	gtctcctttt	gtagccctaa	ttgcttcgcc
88081	tctcatactc	cttggctctg	cgaatctctt	cgcgggcttc	accaatcctc	ctcctctgca

```

79621 aaatgatgta acaggtaaca acaccattgc tattcaagtt gataccagca gttgtacatt
79681 taggagtgtc cttacggtgg ttctctcccg acagctcaac cgaggtaaata acaaatcatc
79741 gtcgtgtgtg tctctctcaa tgactttggc tcaaactcaa ctatgaacta taatcttcat
79801 ctgtacagat tggtttcaaa ggagagtgtg atcactcaga gagtgcacaa cgtatatcac
79861 tgatgatcca aggttctctt cactgactctt gaagtgttga tggtagtttc acagcataac
79921 acgcaacatg aagccacatt tggctcctca gtaggcaaag gcaatgaatg gtagcttccg
79981 aatgtataga tgctttgaga tatcagaaga gaaccccccag aggaacctgt ataaaactct
80041 tctacacgaa gataattaca gagacttgta acttcacttc aatttttggg tttgggattc
80101 ttttttgagt aaaaaggtag gaagagggtta tatttttagt tgtttacttt cttctcatgt
80161 gctctgaaac aaaataagaa acatctccaa cgtattgtag tgattacca ccgataaaaa
80221 gctaaagaat ttttgagtta gtcgattaca aagtataaac acaaatttga attgatacac
80281 caaagatttc ttttaacatt cgtgggaggc cgagtattat ttggctttag cttcgaacca
80341 agactctctg gagatagaag ctctggctaa atacatgacg gactgagagg actacggggg
80401 ctatcttaat ttctttttat atgaattttt ctatgttaat ttaacttttt agatattgat
80461 ttccattgta attaaaacgc tggttgtcct aatttttatt tatatatata ttgtataggt
80521 tccattaac ccacgatcaa aacagtttca tacaactcgc cctcatcttg ttctctaaga
80581 tcaacgtgtc tgatcaagct tgatctgctt ctcttttttt tttgtttttg taattattct
80641 tgcttgattc gatcaacgtg tctgagcttt gattcaagtg ggggtttgga ggtatgtgag
80701 cgtcaaactc tcaagtttta attctggaat taggtttctc ttagtggggt ttttttttgt
80761 atcatccgag tgttgttggg tgaatataaa tataaatoca accactagaa catgttttac
80821 ttactttggg tttgtttttt ctccagatat ataccgccat cgccatggat aatcacttgg
80881 aattagcgat aaaggatgca atcaccgcgg atgatcttaa acgtgtggat caagaaactc
80941 aacacccttt gttagctcag gagcttgatc ttgattcctt ggagaaactc cctcgggcca
81001 caactcatal atatcgggta tactccaaag gtttgggtgag tgaagagctt attaaggatg
81061 atacgatgct ggtcgttggg ctaggtttgt cctctgtga ctcgcacgat tacacgaaac
81121 aggagattaa taaagctctg agaaaccaa agctggcggc acaccagaa gctgcggaac
81181 tggctgccat cattcacggc ttgaaatggg ccttgggaac tggtagcgaa cgtatccaat
81241 tcttctgtga cgactccaat atcttggcct acgtatgttt ctctctcttt cttcttttac
81301 cctgaatcca gtatctgtt tgtacacctc tcatgttgtt ttacaggtt actcgtaaag
81361 ctgcacctaa cgagtcatt gtagcaaaac ttttgagca tgtgtctctt tctcagaca
81421 gattcagctc atgtcaggca cttgcaactg taagcagaga cgactcgtt tctgtcatta
81481 agctagcaaa agatgctata gcttcccaa ctataggtg tgaaggcgac accgagtatg
81541 agagtgtgcc agtctgtctc gcttacgttt cacctaata taagtttgag gtgcaaggct
81601 gcttccaccg catctgcgtt acgtgcatga ggaagccctt ctcatccgaa caaatactac
81661 gagggaaacac agcaatctgc ccttaccggg attgagagaa tgatcttgtg ccagaggatt
81721 gtagagcttt tgctgatgct gatgctatta ctcttatgat ccagcgcaag aaggagaagg
81781 ctatccccgt taaagacaga gctattgtc ccaaccatc ttgttctttt ctgagtctgg
81841 acctcgacct cattaggcac ataagcaaaa atctcggca ctcagaaga gacgagagt
81901 gcattggagt cggcttgtct ttctgcaaaa aatgccatgt tccgtggcac tacaagaaga
81961 catgcgatga gttcaagaag tcggagtctt acctgaaatc tgacgcggcg attttggagt
82021 cttttgtgaa gacacaagga tggaaaaagt gttcccagtg tcagagcatc gttcaacatg
82081 gtggcggctg ccaacaaatg acttgcaggt attggacatt tgttatttat tgagcctact
82141 tttatattat ttgactttaa ctatctgtct ttttgttctt ggcagacatt gcaaacacga
82201 gttctgttac acatgtggcg ctccgtgtaa aaagaagaaa ctgacatgta aatgctcggc
82261 ttcagggaaa taaacatcca aggttttact atacagatca ttccaatag tttcttatat
82321 aacttgtttt gatggtgttg aagatattag tctcttctca tgttttctga aacaacatta
82381 gagacatacc aaccgagaaa cgctaaaact atctcttata aataaagagt ttttgagtta
82441 atcgattaca aagtacaaac acgaatttaa taaaagaata tttgaattga taaacaaat
82501 atgagaggaa gaatacaatg aatcaacgtg tgtggtggct ttaacatttc gtgggaaccc
82561 gagtttttat tggctttaac ttcaaaccaa gactctcttg tgatagaagc tctggcgaaa
82621 tacatgacgg acttagtgga ctccgggggg tatcgctaac ccgaccaggc cgatacaatc
82681 caaaccccat gaaaaagaac tccattgta tcaacattgc atgtggttg atgtgtgtc tctctgtca
82741 ccttctgcat acaagcttgc acctgcaatt caaaaaacat caaagcatgt tccacgaatc
82801 taaaagcctg aatctttagt ctcgagaga agttaagagc taccagagct gtttttacat
82861 taaaccgcat acctcaacca aggcgtgtga ctttctgtcc aatgctgtca ccatgtgaag
82921 agtttcagag tggaaaggag attggtcgcc aacaaagatg agtgtagcgc atttcaaact
82981 cttcaatcca tcagttaagt catgtctcct gttaactgct tctagaaacc gcataagact
83041 actaccgtgt ctttcaccta gcagctgtat aaacatgtcg tatcaacaac gtaatcaagg
83101 tgggtgaaaag cagtaaaatc atctggaaat aactactagg tttaggcggt tagcactcac
83161 tcttctgcat tcatgtacca catccgctc tggaaactca gagctaccac gagcttcttg
83221 tcgtaattta agaacaagat taatcggtta ctattaagta tatacggatc tcattcagaa
83281 atcacaagaa agattacctt actgaagtac ctctgaagga aaatatcttt taacagtcca
83341 gacatgccat agtagtacia taagtttgac acaacctgca ggaagtttta acatcgggtt
83401 atgcgaaaga accttttgta cagtcatatg aactactatg ggaagggtga aaaaagaaac
83461 cttgtaataa aaccattcag accttagagg tgctttgcat agaggcgata taagaataca
83521 acccaaaact cgttctttat gtttaactct caacctcaa acaatcactg ataagtaatg
83581 agttacatta ttgtataatg ttctgagaat gcagctacaa ataaatcctc aagaggcctc
83641 aatgagtaca tacagcaaac aaggaaagga ttaggcacc agctgtgatt cccatgcaca
83701 ttactgcctc gaggtgcaa gagacataga tagttagagg ataatagcgg aaattactct
83761 tgttatgata gcccaatttt agagtggag ttagtcgaaa catgattcaa gatttaccta
83821 aagaagttca atacttcaag aatctggctc gcgaggtctt caacagaagg tgacggatca

```



92401	tgaatgaata	caaaaaaggt	tttttcttgc	attaaatgga	atagaagttt	atggatagaa
92461	tagtttagaa	tttaaagtga	tggaactgta	cattttgtgg	aaaaaacagt	ttgaagtctg
92521	aaccacacaa	ctccttggcc	aagtagcttt	ggaggtagac	tttgctgcat	tcaagtctct
92581	gaacacacaa	tttttattatt	gtattgcata	cactaacttt	tatgtaattt	cagacaaaaa
92641	gatttgtaatt	tttgtcttta	agaccaggat	tgggcccaatc	tttgtttttt	aattttttata
92701	cagccaaaag	taagttctat	t			